Detection of Cognitive Process during Differentiating Task with Electroencephalography

Tan Trong Dang Vo

February 2012

A thesis submitted for the degree of Master of Philosophy
of the Australian National University

For my father and my mother

Declaration

The work in this thesis is my own except where otherwise stated.

Tan Trong Dang Vo

Acknowledgements

The work in this dissertation would not be a possible without the help of the people listed below,

- My foremost appreciation is to both of my supervisors Prof. Tom Gedeon and Dr. Sumudu Mendis. I would like to thank Tom for his kind and calm approach to many issues I faced - and yet still can pursue me to get things done during my time here. I am also thankful for your invaluable ideas and suggestions, as well for the never-give-up attitude. As for Sumudu, being supervised by him only for one year, yet I have learnt a lot from him on how to be a conventional researcher - especially at the beginning of my time here. To me it is invaluable to have both of you as my supervisors as you always there to guide me with just research directions and there to support me.

- I would like to thank my family: my mother, father and sister for supporting me throughout my time here. Studying away from home was not easy for them to support me and I do appreciate your occasional unexpected long drive trip down here to visit me.

- The people in my research group from ANU: Gareth, Dingyun, Nandita for the useful discussions during meetings. Especially Gareth for his insights on various aspects of this research field.

- I also would like to thank all the friends I have made here in ANU for your great companionship and inspiration: Xi, Grace, Richard, Pete, Ting, Nguyen.

- Also, I would like to thanks all members of the Research School of Computer Science at ANU for their help and support in various ways (administrative, technical and so on). My life would be much harder without you.

Abstract

Human beings notice differences in their environment and make explicit and implicit differentiations between events and activities. These are very intuitive activities that humans perform on a regular basis, studying the cognitive processes behind them will help gain more knowledge on mental tasks, opening up possibilities for better brain-computer interaction (BCI) mechanisms in the future. In this study, these cognitive processes were measured using a non-invasive brain measuring method, Electroencephalography (EEG) as well as a behavioural measure (eye gaze path). Separate experiments were carried out to examine both scenarios of explicit (voluntary) and implicit (involuntary) differentiations. Using a combination of machine learning tools for signal analysis, a high correct rate of classification results in combination with significant statistical figures achieved has shown that one can effectively identify those mental tasks from studying the brainwave and eye gaze measurements.

Organisation of the Dissertation

In **Chapter 1**, we (you, the reader and I) will discuss the background knowledge that is the foundation of this research work. The chapter starts by introducing the fundamental components that my research is based on: Brain-Computer Interface (BCI) and Electroencephalography(EEG). This goes as far as discussing two conventional BCI methods, Visual Evoked Potential and Mental Tasks that represent two different approaches in designing BCI systems. They are related to our research because they are already proven BCI systems and, to some extent, provide mechanisms of communicating commands to a computer via EEG. The chapter finishes with a discussion on fine-textured differentiating actions in the context of BCI, which will put into some perspective about Chapter 2.

Chapter 2 of this dissertation is dedicated for describing our work on detecting explicit differentiating actions from EEG signals in detail. We start this chapter by examining the challenges that we could encounter as well as investigating the future potential of the work. Subsequently, we introduced considerations taken into account for the purpose of designing suitable EEG experiments to achieve the research goal. The chapter then moves on to describe the experiment in terms of set-up, hardware and software options, methods in signal processing and data classification. The chapter also explains in detail the two machine learning techniques employed for the purpose of classifying experimental data: Support Vector Machines and Artificial Neural Networks. The analysis section of the chapter summarises the results in various aspects and is followed by a discussion of the future work that could be done as an extension to this work.

Chapter 3 describes our investigation on the ability to detect the existence of implicit differencing activities from studying EEG signals. For this purpose, we carried out reading experiments to confirm our assumptions: by studying a person's EEG signals, we can pick up his/her unintentional differentiating activities while reading English paragraphs. For that reason, a portion of the chapter is dedicated for explaining the experiment's background while the rest of it describes the experiment in term of set-up and the methods by which the EEG data is processed. Finally, the chapter concludes by detailing the classification results from the two aforementioned classification techniques and summaries the outcomes of this work.

Chapter 4 describes further research on human behaviour in reading tasks. Unlike the general theme set so far by this dissertation, gaze tracking was used as the mean to capture biological features for studying. Our aim is to investigate if we can associate certain gaze behaviours to reading tasks. Our hypothesis is that if we could capture a person's gaze points while he is reading, we would be able to either tell what subjects he/she is interested in, how quickly or thoroughly he/she can read or simply if he/she is paying attention to certain screen areas. The chapter also proposes a novel method of detecting the level of engagement in reading based on a person's gaze-pattern. In the conclusion of the chapter, we compare the performance of data classification using gaze tracking with EEG. We can then confirm the correlation between the two.

In the **concluding chapter**, we summarise the results of the research work as a whole, reflect on lessons learnt and make some suggestions for future work.

List of Publications

My contributions have been published in internationally refereed conferences as listed below:

1. Tan Vo, B. Mendis, and Tom Gedeon. Gaze Pattern and Reading Comprehension. *Neural Information Processing. Models and Applications - 17th International Conference, ICONIP 2010.*

2. Tan Vo, and Tom Gedeon. Reading Your Mind: EEG during Reading Task. *Neural Information Processing. Models and Applications - 18th International Conference, ICONIP 2011.*

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

1.1.1 Differentiation Tasks

Within the context of this research, only differentiating activity of human as the act of perceiving the difference in or between objects was considered. In other word, it is to mentally discriminate one thing from the others, a very common activity that humans perform on regular basis.

In general, this mental differentiating process is a process that may have been studied over the years. However, this very topic may have been studied under different name or contexts. Up until now, various *theories* have been introduced to explain how people *differentiate* objects they observe, as well as identify influencing factors. It is however, still a very challenging research area as there are a lot of unknown *variables* involved.

One could broadly categorise differentiation activities into two types according to the way a person perform them:

- **Explicit differentiation** Refers to the voluntary discriminating actions that are performed by human when provided with "choices" to decide between.
- **Implicit differentiation** On the other hand, study of implicit differentiation deals with hidden discriminations/ judgements made by a person without their acknowledgement.

Because of that, the study of this tasks could be devided into two sections: one for investigating the nature of explicit differentiation while the other deals with implicit differentiation. The goal is to see if one can computationally recognise those mental tasks from studying EEG signals captured from test participating.

So how do one to achieve this goal? Electroencephalography (EEG) was used to measure the brain activities from electrodes placed on the human scalp and attempt to determine various cognitive processes by studying those EEG measurements (signals). In this scenario, a black-box approach was considered: If one could feed accurate and steady enough inputs to a person's brain, he could potentially be able to observe certain outputs from the person's EEG signals captured during scenarios. The inputs in this context are stimuli sent to the person that result in cognitive reactions from his brain. Two different approaches for both of types of differentiation are required in regard to experiment designs.

With regards to studies of Explicit differentiating tasks, it would be desirable to design experimental scenarios that guides a person (equipped with EEG) into making various differentiation tasks and expect those EEG captures to exhibit some patterns that can be recognised by machine (i.e. with the use of machine learning algorithms and so on).

On the other hand, a more subtle approach is needed for designing experiments on Implicit differentiation actions. In this dissertation, Reading task was used as the mechanism to conceal the differencing tasks from the knowledge of the person (hence implicit). The selection of reading because of its familiarity to most people, also the complexity of the activity will help distract the person from recognising any discrimination he makes about the contents he is reading.

The goal of this work is to confirm EEG's potential in detecting series of differentiations made by humans while performing another mental activity such as reading.

Studies of human brain activities have generally been done within the perimeter of either neuroscience or psychology research fields. This research, however, belongs to a relatively new trend of research on human brain that takes advantage of modern computing power to aid to the study of human brain.

1.1.2 Brain-Computer Interface

Traditionally, humans interact with computers via peripheral devices such as keyboards and mice. Mice and keyboards are not necessarily the most natural interaction methods with computers and, in some circumstances, may not be capable enough. There are also other input methods that are considered *alternatives* to keyboards and mice. Those are methods that utilise *eye-gazes* and *voice*. However, the prospect of them being used as conventional peripherals is well in the future.

Brain-Computer Interface (BCI) is, in a sense, an alternative human-computer interaction method that involves a *bi-directional* link between human brain and computer. Being Bi-directional in the sense that a person's brain issues *commands*, *instructions* to a machine and the machine provide *feedbacks* (visual, auditory) to the person - not directly to his brain. In some cases, such as for people who have severe damage to their spinal cord, BCI is not just a *convenience* but a *necessity*. Think about those with a severed spinal injury and cannot use computers with their limbs, BCI is *technically* a very suitable solution for them. According to Lebedev and Nicolelis [35], in general, BCI systems can be classified as either *Invasive* or *Non-Invasive*:

Invasive BCI Electrodes are implanted within the cranium of the brain. By BCI standards, this methodology provides neural signals of the best quality and has a high potential for further development. However, its biggest drawback is that it carries *risks* associated with *neurosurgery*. Neurosurgery also greatly limits this technique's accessibility for general use.

Non-invasive BCI These methods *do not* require any invasive surgical procedure. Popular examples of these are the ones that based on magnetic resonance of blood vessels to identify brain activities (brain-mapping) - fRMI for instance. On the other hand, techniques such as Electroencephalography (EEG) utilise electrodes to study the brain activities from the outside of the human head (normally over the scalp area). Being without the need for neurosurgery, these methods are considered more accessible than their counterpart. Nowadays, some EEG devices are sold as consumer products.

From this point of the dissertation onward, the only type of BCI will be referred to is non-invasive.

1.1.3 Electroencephalography

As for the present, the use of EEG readings as a non-invasive technique for Brain-Computer Interaction (BCI) is a popular choice among research institutes as it is considered safer and more accessible than the aforementioned invasive (direct brain control) methods. EEG fundamentally is the recording of the electrical activity of the brain from the scalp with electrical electrodes. The recorded signals reflect the electrical activities on the surface of the brain, which is influenced by the electrical activities from the brain structures underneath the cortex. These signals are measured in microvolts (μV) due to their *small* nature.

The first recordings on human were made by Hans Berger in 1929 although similar studies had been carried out in animals as early as 1870 (Henry [27]). Since then, EEG has been involved, developed and still being used widely for clinical purposes i.e. evaluation of brain disorders; yet there exists a new trend in research that utilise EEG as means to interact with machines. That is also the focus of this research work: making use of EEG signals from the HCI perspective.

The studies of EEG signals generally required *transforming* EEG channels into *frequency bands*. The amplitudes (or combination of) belong to those frequency bands are normally used as *features* to study with [27]:

Delta band Less or equals to 3 Hz.

Theta band Between 3.5 Hz to 7.5 Hz.

Alpha band Between 7.5 and 13 Hz.

Beta band Between 14 Hz to 30 Hz (beyond 30 Hz is generally considered as noises).

In term of amplitude, there are brain activities identified to be associated with the activations of an individual band, or a combination of more than one bands. However those *activities* will not be discussed here as for this research, the interest in studying the activations of those bands *purely* on the statistical basis; it was not planned to explain them in any other perspective. There are also other methods to study EEG such as those that study voltage, morphology, synchrony or periodicity or EEG waveforms. They also will not be discussed in this dissertation as EEG signals only been analysed within the *frequency domain*.

In regards to that, generally, there are two broad categorisations for EEG-based BCI systems [21]:

Synchronous BCI driven by the brain event-related response to *external stimuli*, normally provided by the BCI system's GUI. The most well known example of this types is the *Visual Evoked Potential* (VEP) system. The synchronous approach has certain successes in recent years in terms of accuracy and the information data throughput.

Asynchronous BCI driven by the *voluntary* modulation of the brain activities. These systems do not depend on brain reaction to the stimuli; instead, they depend on detection of *steady*, *pre-defined* brain activities to drive the system. One example of these is a system that based on *mental tasks*, which requires its user to perform a set of mental tasks repeatedly to indicate his commands

In the next two sections 1.2 and 1.3, *VEP* and *Mental Tasks* will be discussed to demonstrate the differences in approaches that utilise EEG in research. Also, from this point of the dissertation onward, the only type of BCI systems referred to is the BCI systems that utilise *EEG* as the method for measuring the brain activities.

1.2 EEG and Visual Evoked Potential

1.2.1 VEP Fundamentals

VEP fundamentally depends on a distinct phase of cortical processing that can be captured from the cerebral surface by EEG at the very moment a response occurs to a visual stimulus occurs on the user interface. A simple system that could demonstrate VEP would be:

- A menu-based interface is given to the user on a computer screen.
- One after another, each menu item will repeatedly flash for about 0.5 seconds.
- To *select* an item of the menu, user just need to focus on that item and wait for the item to flash.
- At the time the focused item flashes, the EEG can capture the neurological responses from the user's brain and will consider the user *has selected* the item.

Responses are recorded from electrodes that are placed on the back of the head and are observed as a reading on an electroencephalogram (EEG). The reason for that is because the visual cortex, the area of the brain that is responsible for receiving and processing visual information, is located just above one's neck towards the back. Hence, one would expect certain brain activities to happen around that area when a person receives a visual stimulus. There are certain benefits of VEP system over different types of EEG-based interface systems:

- Input accuracy is generally high. In some special cases, for instance the brain activities used as the binary input in experiment by Ferreira at el. [19], the accuracy can reach 94%. Galn et al. [21], who developed a new VEP-based wheelchair claims an approximate 65% accuracy for the system.
- Relative ease of training effort required from users. The nature of the system does not require users to produce intense and exhausting level of concentration required by other approaches.

1.2.2 Current Research in VEP

In the early 2000s, researches on visual evoked responses reported that brain reaction to visual stimulus is actually a combination of EEG components that can be recorded from different brain areas (apart from the area around the visual cortex). The reports from Makeig [40], Smith [30] and Hoson [28] have identified such components. Detailed findings are as follow: Mazaheri [7] and Hanslmayr [26] on the Alpha component and Edwards [17] on the Gamma component.

Odom et al. [44] proposed guidelines for VEP experiments and practices. The paper pays attention to the EEG procedures and the interface aspect of VEP. The paper however does not cover some specialized VEP types and also leaves the detailed analysis of the response patterns as an open topic.

Kimuchi [32] and Klistorner [33] took a more detailed look at the use of multifocal VEP(mVEP). The outcomes of these papers indicate that the results of VEP are highly dependent to the electrode locations. Given the findings of Ian [12] on the topographic EEG mapping to local brain functions, one could assume that left-right brain nature of test subjects could affect the VEP test. This is interesting because these are involuntary activities of the brain, which are known not to be influenced much by the left-right brain nature of the test subjects.

Positive 300 wave is a positive deflection in voltage at latency of roughly 300 ms after the point of visual stimulus occurs (Figure 1.1). The reason P300 is mentioned in this paper is because most of the implementation of a VEP system base around the activities of P300. Salil and Pierre [45] have reported their finding on P300 and its relation to Negative 200 wave, which happens right before it. Until recently, research groups have used P300 to develop EEG-based interface for their projects. From a simplified menu-driven command system [46] to a more sophisticated system where stimulus are placed on a representation of a real world navigation scenarios [21]. Other possible usages of P300 with VEP are more complicated input devices such as virtual keyboards.

1.2.3 Drawbacks of VEP

The main drawback of VEP is the relatively low information transfer rate [21][46]. There are quite a few approaches to address this drawback:

- Eliminate unnecessary visual cue (noise) on the interface. It is essentially to reduce the failure rate. But one can compensate that with a better recognition algorithm

- Optimise the visual stimulate cue patterns. One possible way is trying to predict what the user would like to choose next in the context and then subsequently increase the frequency of the visual clues appearing at the predicted targets.

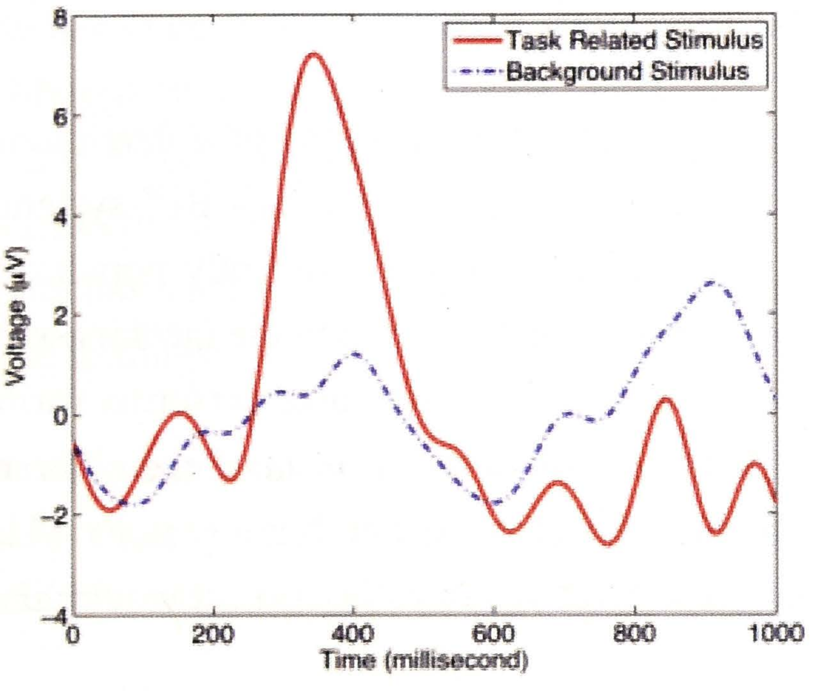


Figure 1.1: Multi-trial averaged bandpass filtered P300 signal from one electrode (Task versus Non-task stimulus cases)

Another disadvantage of VEP is that it still inherits the non-stationary nature of EEG-based system. This means that complicated calibration process and low accuracy in the long run are still inevitable.

1.2.4 Advantages and Opportunities

A recent demonstration of a user using VEP to use *twitter* on a virtual keyboard with a reasonable input rate [1]. It is suspected that he could have used a built-in dictionary to predict what *letter* would likely to be chosen next, hence the system could allow him to *select* the next character faster. Other opportunities that one could explore with VEP are to experiment with different way for laying out the commands/choices on the interface so that:

- Users could *select* an item by fully concentrate on the chosen item.
- Users could *select* more than one items or a group of related items.
- Users could *interact* with a dynamic interface rather than being driven by it.

Those proposed ideas above could be used in conjunction with other HCI techniques like eye-gaze or head-tracking to make them possible.

1.3 EEG and Mental Tasks

Being research that focuses on the EEG patterns of *differentiating* tasks, it would be valuable to investigate another EEG research area - BCI systems driven by the brain activities performing *mental tasks*. This is a recently popular class of BCI solution used for commuting commands from the brain to the machine-side of the system. During an EEG *Mental task* session, the system user performs various mental tasks such as *counting*, *rotating* imaginary objects, etc *in his head*. These mental tasks, being *selected* based on their *distinctive* activities in brain regions [41], are then associated with various designated actions/commands the subject would like to perform throughout the sessions.

1.3.1 Background

Utilising mental tasks in BCI has been referred to as a clever trick to enable BCI systems, with minimal effort, interpret a set of multiple commands based on brain signals. Designated mental tasks are chosen so that if one *performs* them inside his head, they will triggers brain activities in *different combinations* of brain regions. Those combination patterns can then be recorded by brainwave recording devices, then "translated" into commands that are assigned to them. The study of mental tasks is known as Cognitive Task Analysis (CTA). This research area is generally carried out by psychology research groups. In fact, the research methods employed are categorised into three broad categories [47]:

1. Observation and interviews

2. Process tracing

3. Conceptual techniques

Those above are also essentially psychological research methods in studying such tasks. Understanding the complexity of a task in term of *mental load* is the key factor to determine the potential *task candidates* of the research. A more recent research movement in this area is the use of *neuroimaging* techniques such as *Magnetic resonance imaging (RMI)* or *Functional magnetic resonance imaging (fRMI)* to aid the study of brain activities. (fRMI is the preferred method out of the two). These devices offer excellent spatial resolution (can be as good as 1mm accuracy), making them the ideal choices for observing activities in *brain regions* while performing different mental tasks. That movement has tied the bond between CTA and BCI researches tighter than ever. Detailed studies of each mental task by studying fRMI capturing are in fact very relevant to EEG-based BCI systems. Unlike fRMI, EEG is biased towards the cortical surface, results in a very low spatial resolution compared to fRMI's brain imagery. So where does EEG fit into the picture? It would be toward the end of the chain. fRMI helps studying the brain activities in regions for *mental tasks* and researchers in EEG benefit from those findings:

- Tasks selection: choosing tasks based on the brain activities that would yield the best result. A task requires a left cerebral hemisphere lobe activity vs. a task that requires right cerebral hemisphere activities, for instance.
- Signal processing: optimising *signal processing* components to fine-tune the system in detecting the *expected* mental tasks.

With the raw EEG channels, it is very difficult to clearly separate signals from each individual channels. It, however, is still *practical* enough to identify brain activities (hot spots) on different areas of the brain, giving there are enough *repetitions* of mental task activities (a very common practice in dealing with *noisy signals* like EEG captures). With the appropriate use of filters, data cleansing techniques, the accuracy of recognising *mental tasks* could range from medium to high. Furthermore, EEG offers better *temporal resolution* compared to fRMI as well as EEG equipment is much more accessible than fRMI equipments. The combination of those has made *EEG* and *Mental tasks* a very feasible solution for realtime BCI systems.

1.3.2 Success Story

- This is just one example of this type of achievement. Craig [14] and his team designed a brain-controlled wheelchair using this approach: utilising mental tasks such as *Mental Arithmetic*, *Figure Rotation*, *Mental Counting* and *Letter Composition* to map to *navigating instructions* of the wheelchair. According to Craig [14], one can effectively recognise those four complex brain patterns in the *frequency domain.*

- EPOCH Emotiv: one of the very first fully capable EEG devices that is being sold to the consumer market. The device SDK has the ability to study mental tasks such as *Push*, *Pull*, *Lift* and then allows users to *map* them into actions of their applications, such as games. On Jul 2010, Tan Le, head of Emotiv Systems, demonstrated this device's capability on TED [6].

1.3.3 Discussion

In summary, this area of the research is well understood and closely followed by psychological researchers (Iain [38]). Many of its success in recent years is due to the use of neuroimaging technologies such as fRMI method, and since then, has provided a wealth of information on cognitive activities behind numerous mental tasks. fRMI, with a relatively good spatial resolution is a great candidate to be used with CTA based systems. EEG, despite its significantly lower spatial resolution than fRMI, is still a very popular choice to be used with CTA because of its far more superior temporal resolution and accessibility compared to the other methods. Mental tasks and EEG together produce potentials in both *practical applications* and *theoretical research*. However, at the end of the day, it is still a mechanism for transferring *commands* to machine by making use of differences in brain activities between mental tasks. In other words, the focus of these researches was never to be on the nature of mental tasks in the context of BCI. However, the background study of this research area provides the idea of analysing the EEG patterns on the frequency domain rather than in time domain. If one could model the mental task in frequency, modelling the brain activities during *differentiating* tasks would be quite plausible.

1.4 Differencing Task and EEG

1.4.1 Background

In practice, a *differentiating* process is a process that *may* have been studied over the years. The use of the word *may* here because this have been studied in different contexts or to serve different purposes that are not related to this work. Up until now, various *theories* have been introduced to explain how people *differentiate* objects they observe, as well as identify the factors that influence it. It is however, still a very challenging research area as there are a lot of unknown *variables* involved.

A person, given the task of differentiating two or more objects, will eventually reach a definite conclusion/choice. However, the process of *differentiation* that leads to the *conclusion* has already completed by then. The process finishes in a very quick, almost instantaneous, manner. He normally perceive the *conclusions*, but he rarely recognises the processes that lead to them. This would apply regardless of whether it is of implicit type or not. Within the BCI perspectives, however, studying the *outcomes* of such conclusions would be quite *impracticable*; it would be more plausible to shift the focus to the *differentiating* processes associated with them instead.

It is anticipated that these processes, even under different circumstances, still share something in common that can be detected from analysing captured EEG signals. According to Vanrullen and Simon [50], the mental differentiating process starts with a perceptual, task-independent process followed by another task-related, category-independent process. So under controlled experiment condition, in which one are expecting to extract out those *precious EEG patterns*, the whole process normally just takes about 75-80 milliseconds [50], provided the differentiating task is fairly simple. If one could design an experiment, in which he:

- Guide a test participant to demonstrate all of these above cognitive processes.
- Capture his brain activities.
- Repeat the above steps, as many times as needed so that it's possible accumulate a dataset with sufficient data in it.
- Study the dataset.

If the results are *positive* and *statistically significant* enough, one could then claim that he has identified the cognitive process of differentiating via the use of EEG.

1.4.2 Research Goals

The main goal of this research is to focus on using machine learning techniques in detecting the mental process happening during human differentiation especially with machine learning tools such as *Artificial Neural Networks* and *Support Vector Machines*. Having considered various alternative approaches, the problem was tackled by overlooking certain relevant background knowledge, such as mental tasks theories or perspectives of neuroscience on the topic. Instead, sensible assumptions were made in places while just focusing solely on reaching the goals. The main goal is to answer the following questions:

- In regard to differentiation, is there a *pattern* of EEG signals that exists for *specific* types of differentiation? (Individual case)
- In regard to differentiation, is there a *pattern* of EEG signals that exists across *different* types of differentiation? (General case)
- Are there any interesting and distinctive features of EEG signals that one could benefit from?

1.4.3 Considerations

In reality, the nature of EEG and the complexity of almost every human cognitive process make the above assumption somewhat naive. These are the briefs of significant challenges that researchers doing similar work have to face:

Diversity in Types of Differentiation

A human has to make differentiation at all levels on a daily basis. It is so much that establishing a categorisation system for those is a challenge all by itself. For instance, could one hypothetically take *difficulty level* as a category for classifying the differentiation types? If *instinctual reactions* were considered as the *least difficult* type of differentiation for a person to make, how would one then define instinctual? How would one go and design an experiment to produce that? On the other hand, if the most contradictory paradox was chosen as a guideline for the *most difficult* differentiation, the process to get to the conclusion may take *too long* and it would involve quite a few *undetermined* variables. Could one afford to capture and study the EEG signals that associate with it? And that is only the attempt to categorise *differentiation* based on *difficulty*. Hence one can choose some features/attributes for the purpose of categorising *differentiation*, but would not be expected to come up with a *definitive scheme* for that purpose.

Influential Factors

There are several factors that influence this very intricate process. Some of these factors were briefly listed by Cindy [16] as past experience, cognitive biases, age and individual differences, belief in personal relevance, levels of commitment, influence on the choices that people have to make. The importance of understanding those factors could help to know the differentiating process better, as they effectively shape the outcome of this differentiation. However, the very similar problem will be faced in terms of comprehensively categorising and understanding such factors.

Existing Effort in Modelling Differentiating Process

Everyone in different fields that is interested in studying the differentiating process will likely to have his very *own* model for it. Such models in philosophy, psychology, economics, statistics and even mathematics are normally concerned with the values, uncertainties and other issues that are specifically relevant to the corresponding field. The issue arises when one need to study *differentiating process*, suddenly finds this situation when these models are overlapping, or even contradict each other. This is due to the nature of the problem domains, which generally span across multiple research fields. Because of that, designing a BCI system according to existing models is a sizable challenge for BCI researchers.

The Time versus Complexity Factor

This is one of the more relevant challenges to this research. As one can only capture brainwave activities as fast as the computing power allows, the resulting data can never be as *realtime* as he would like it to be. Unfortunately, differentiating processes can be very quick (in term of factions of a second), there is a good chance that he could not fully capture, or in the worst case, miss them completely. Designing the scenarios so that test participants have to take longer to make a differentiation would introduce more variables and complexity in data analysis. This challenge is indeed the most difficult one to properly address.

Chapter 2

Explicit Differentiation and EEG

Throughout the day, a typical person would be overwhelmed with activities of differentiating different things and at various scales. A differentiation is considered as explicit if the users are provided with the options forthrightly (visually or auditory etc.) and that he voluntary makes the distinction between them. Making such differentiation, little or big, easy or hard, still requires combinations of cognitive processes occur across sections of the brain. It is believed that studying the underlying mental processes happen during those activities could lead one to more knowledge on the properties/features of the differentiation that have just been made. And to go one step further, one could be able to tell the outcomes of them (conclusions). An EEG based BCI experiment was organised to study the nature of these processes. The result obtained was strong enough to confirm that it is possible to computationally detect explicit differentiation activities from the EEG signals.

From this point onward till the end of chapter, any reference to *differentiation* will be considered as of the *explicit* type only.

2.1 Introduction

2.1.1 Research Goals

The main goal of this research is to focus on using machine learning techniques in detecting the mental process happening during human differentiation especially with machine learning tools such as *Artificial Neural Networks* and *Support Vector Machines*. Having considered various alternative approaches, it is decided to tackle the problem by overlooking certain relevant background knowledge, such as mental tasks theories or perspectives of neuroscience on the topic. Instead, sensible assumptions were made in various places. The main goal is to answer the following questions:

- In regard to differentiation, is there a *pattern* of EEG signals that exists for *specific* types of differentiation? (Individual case)
- In regard to differentiation, is there a *pattern* of EEG signals that exists across *different* types of differentiation? (General case)
- Could the patterns, if existed, be consistently found on multiple test subjects as well as on each individual?

An EEG-based BCI experiment will be desinged to helps achieving that research goal. It will require efforts in areas such as *signal processing*, *pattern learning* and *classification* of EEG data.

2.1.2 Methods

The main goal of this research is to focus on using machine learning techniques in detecting the mental process happening during human differentiation especially with machine learning tools such as *Artificial Neural Networks* and *Support Vector Machines*. Having considered various alternative approaches, the problem was tackled by overlooking certain relevant background knowledge, such as mental tasks theories or perspectives of neuroscience on the topic. Instead, sensible assumptions were made in places while just focusing solely on reaching the goals.

2.1.3 Brain-Computer Interface

The fundamental principle of BCI systems is the *interaction* between the *human brain* and *computers*. In practice, a typical BCI system would comprise of more than one component to be able to fulfill that purpose successfully. These *building blocks* of the BCI system are shown in Figure 2.1 :
The *interactions* in this context can be described as either one-way (brain $\rightarrow$ computer) or two-way (brain $\leftrightarrow$ computer), which depends on how important the *feedback* loop is to the system. For *synchronous* BCI systems, the system *feedback* is critical as the user has to be constantly fed *stimuli* from the computer and then conveys his differentiating action based on his *reactions* to them. Such systems are the types that take the advantage of Visual Evoked Potential and P300 signals. For asynchronous BCI systems, that *feedback* is not as important (hence the interaction is one-way). Ideally, the user of

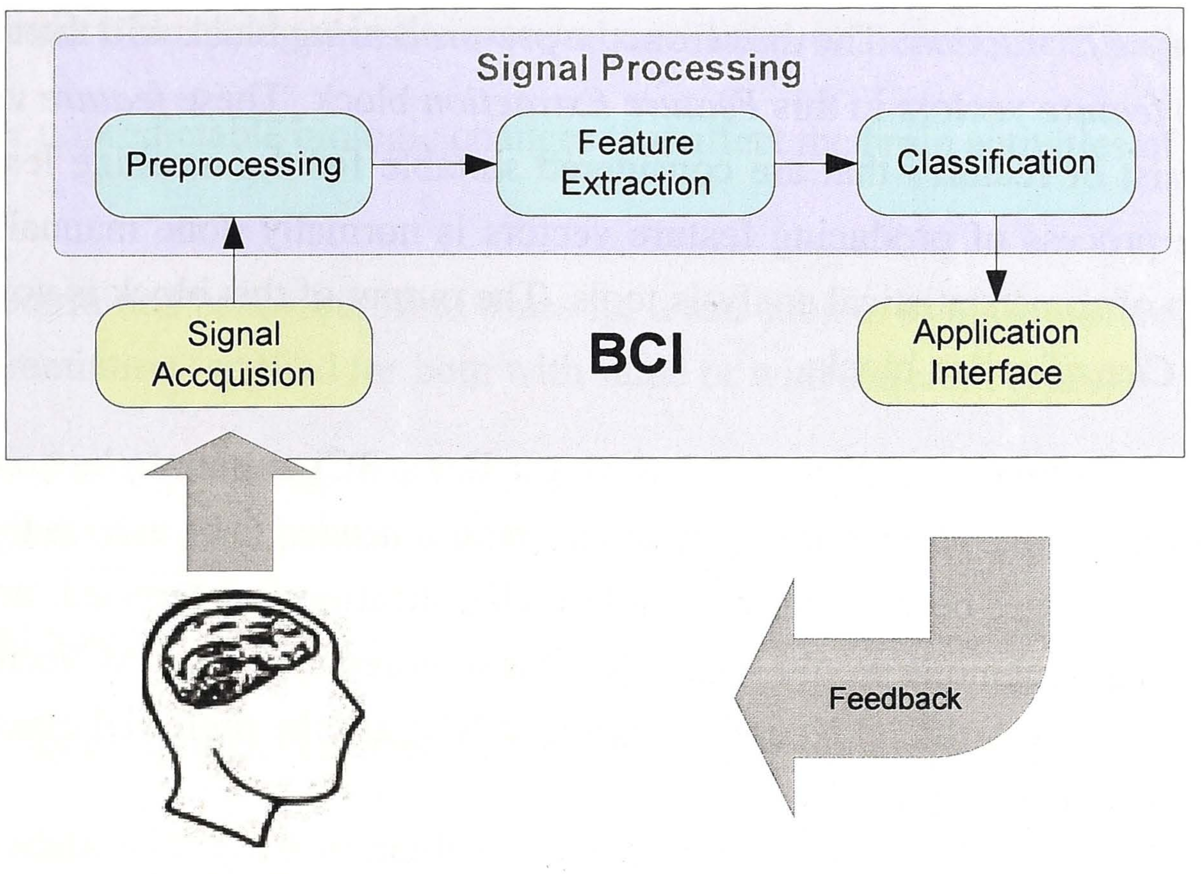


Figure 2.1: Brain-Computer Interface Building Blocks

the system would actively and voluntarily *feed* commands/instructions to the computer by his brain activities that can be captured with EEG. The feedback from the PC is mostly to confirm the issued instruction. In general, asynchronous systems are harder to construct properly as EEG signals are too noisy and unpredictable to deliver robust and timely enough interactions. The system to be built in this research is considered to be among these types of BCI system.

The explanations of the remaining *building blocks*(Figure 2.1) will be related to this specific project:

The *Signal Acquisition* block in this chapter basically covers the *Hardware* aspect of the project. EEG raw signal are captured using either:

- ActiveTwo system from Biosemi: A medical-graded EEG system. It currently supports capturing with 16 EEG channels.
- EPOC EEG headset from Emotiv: One of the new consumer BCI devices. The system records with 14 EEG channels.

As also from Figure 2.1, these following building blocks are generally done with *Software*:

- *Pre-processing*: EEG signals will be band-passed for noise reduction and then be transformed into frequency domain to be ready for the next block. The Fast Fourier Transform (FFT) algorithm will be used for this transformation duty.

- *Feature Extraction*: The data from the *pre-processing* block will then be reduced into *feature vectors* in this *Feature Extraction* block. These *feature vectors* only consist of features that are considered suitable for the machine learning task. The process of producing feature vectors is normally done manually with the help of some statistical analysis tools. The output of this block is now ready for the *Classification* block.

- *Classification (with Machine Learning)*: For a BCI system to be considered *effective*, a *supervised* learning method is recommended to be used as the classifier. At this stage, because only two-class classification is interested, any machine learning technique would suffice for this requirement. Support Vector Machine (SVM) and Artificial Neural Network (ANN) are the preferred classifiers to be used in this project.

- Any visual feedback will be given back to the user by the *Application Interface* block.

The two most important blocks of this BCI experiment are the *Pre-processing* and the *Classification* blocks. The explanations for each of these blocks will be discussed in more detail later into the chapter.

2.1.4 Discussion on EEG

Introduction

Many bio-signals contain a significant random component (stochastic), especially those which are produced by the averaging of many signal sources like EEG. Identify and address the *uncertainties* found in the EEG signals from the brain's processes during differentiating is indeed a challenging task. Furthermore, EEG, being electrodes "hooked up" to human brain via the scalp, will also inherit factors that will greatly reduce the fidelity of the captured signals compared to the original sources.

Influential Factors

This is the list of factors that most likely to affect the recording and the study of EEG signals from test subjects:

- Unpredictable changes at the *stress level* of test subjects.

- Unpredictable changes at the *concentration level* of a test subject.

- Unpredictable changes in the external environment around a test subject.
- Other unpredictable biologic changes that affect the brain activities of a test subject.
- Undetermined behaviours and times for a test subject to make different types of differentiation (applied for both with same or with different types).

Apart from these, capturing brain activities by EEG signals also has the following *undesired* attributes that attention should be paid to:

- Signal noise.
- Inaccurate captures because of improper equipment preparation on test subjects.
- Degradation in signal strength as the experiment progresses (conducting gel getting dried up, for instance).
- Unpredictable changes in signal strength because of equipment malfunction (electrodes get worn out, for instance)

The usual approach of signal processing EEG data is to performed effective noise filtering (band-pass, detrending, etc.). And if it was not enough, there are extensive range of signal processing tools that could be used in this area that could be employed to aid the researchers: Independent Component Analysis, Blind Signal Separation, or simply other types of filters (static / adaptive) that may raise better results.

Discussion

The difficulties in identify and eliminate the non-deterministic components of the captured signals are the main reason why most of EEG-based BCI systems are clunky and impractical in real life usage.

EEG signals, being the *summation* of multiple electrical activities captured from human scalp, do indeed heavily suffer from *signal noise*. To deal with this problem, researchers normally utilise statistical techniques such as *blind source separation* or methods that are based on *Gaussian Distribution* to select the *significant components* over the other components. These significant components are candidates considered to be used as the *features* in pattern recognition/classifying step. The above point implies that EEG-based BCI systems require *calibration sessions*. These sessions normally involve EEG researchers capturing short EEG sessions that cover *all* possible operating scenarios of the system for offline study. They are considered inevitable and are prerequisites to any online BCI operation. An indication of a successful BCI system would be a correct balance between the efforts put into calibration and the effectiveness of the produced feature vectors afterward. Another way to reduce the significance of these problems is to address them at the *signal processing* step. For non-stationary signal processing, adaptive filtering [49],[18] is one of the solutions. Adaptive filtering does employ cost functions to adjust its coefficients (weights) *during* signal processing. The goal is to *progressively* obtain more accurate reduction of irrelevant components in source signals. The changes in the filtering parameters are also accordant to the current state of signal capturing (hence adaptive). If time allows, the performance of popular adaptive filter methods such as Kalman filter [31] against EEG signals should be evaluated. The ideal outcome of this work is to see if adaptive filtering can help addressing the reduction in signal quality as the EEG session prolongs, when the conductivity of electrodes placed on the scalp degrades.

2.1.5 Support Vector Machine (SVM)

Introduction

This this work, Support Vector Machine(SVM) will be used to classify processed EEG signals. In general, SVM is considered to be a *margin* classifier because its goal is to identify the maximum-margin hyperplane that can separate the data points during the training stage. This hyperplane is defined so that it has the largest separation (hence maximum-margin) between the data points belong to two classes. Although SVM can also be configured to be used a multi-class classifier, only two-class SVM was focused on in this dissertation. According to Boser [9], the formula for the output of a linear SVM is as followed:

$$f(x) = w \cdot x + b \tag{2.1}$$

$f(x) = 0$ is the separating plane(hyperplane) between the two class. The *support vectors* are the data points that lie on the two margins from the hyperplane. w is the normal vector to the hyperplane and x is the input vector. This dissertation will not dwell into the details of SVM *optimisation* problems. Instead discussion on the*Non-linearity* and *kernel method* aspects of SVM will be skipped.

Kernel Functions

Non-linearity: If the data could not be *linearly* separated, one could projected them into an higher dimensional space where the data points effectively become linearly

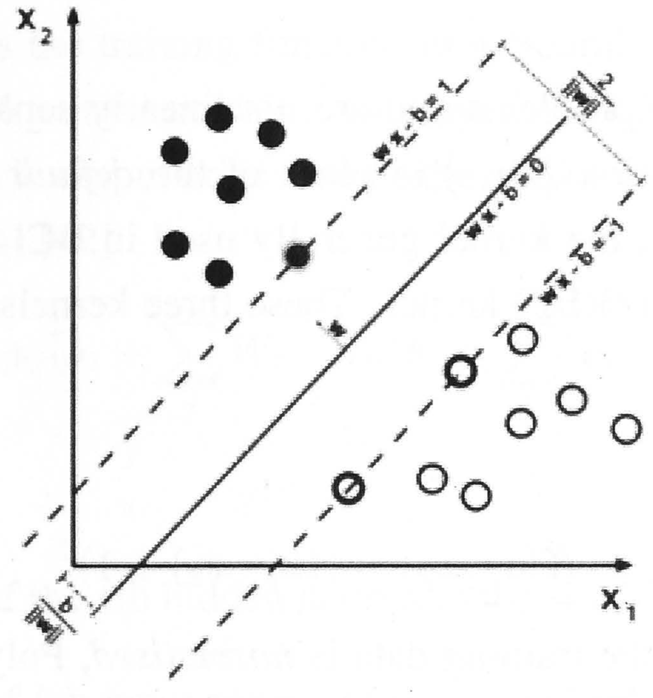


Figure 2.2: Support Vector Machine with hyperplane

separable:

$$x \mapsto \Phi(x) \tag{2.2}$$

Re-applying the new feature vectors back to the original output function:

$$f(x) = w \cdot \Phi(x) + b \tag{2.3}$$

Any *linear* learning machine can be extended to function in nonlinear in input space X by explicitly transforming the data into a feature space Φ using a map $K : X \mapsto \Phi$. SVM, however, can achieve that so *implicitly* thanks to its nature: it only relies on *dot products* of input feature vectors X. These projections are regarded as *kernel methods*.

$$w = \sum_i \alpha_i Phi(x_i) \tag{2.4}$$

Instead of w, one can optimise α_i, the weight of training example, as followed:

$$w = sum_i^m \alpha_i \Phi(x_i) \tag{2.5}$$

Hence, the SVM *differentiation function* can then be written as:

$$\begin{aligned} f(x) &= \sum_i \alpha_i \Phi(x_i) \cdot \Phi(x) + b \\ &= \sum_i \alpha_i K(x_i, x) + b \end{aligned} \tag{2.6}$$

Where K is a dot product function. K is in fact the *default* linear kernel employed by SVM:

$$K(x_1, x_2) = (x_1 \cdot x_2) \tag{2.7}$$

In regards to classifying problems that are not linearly separable, one can substitute a non-linear *kernel function* $K(\cdot,\cdot)$ in place of the *default* dot product K function. According to Lotte [37], the kernel generally used in BCI research is the Gaussian or Radial Basis Function (RBF) kernel. These three kernels were considered for this work:

Polynomial Function

$$K(x_1, x_2) = ((x_1 \cdot x_2) + 1)^d \tag{2.8}$$

For problems where all the training data is *normalised*, Polynomial kernel is recommended. The polynomial order d allows one to customise the feature conjunctions.

Radial Basis Function[29]

$$K(x_1, x_2) = exp(\frac{-\left\|x_1 - x_2\right\|^2}{2\sigma^2}) \tag{2.9}$$

The σ parameter is adjustable. It essentially defines the width of the RBF curve and would dictate the performance of the kernel. This parameter should be carefully tuned to individual problems.

Multilayer Perceptron (MLP) kernel:

$$K(x_1, x_2) = tanh(\kappa x_1 \cdot x_2 + c) \tag{2.10}$$

Also known as the Sigmoid kernel. This kernel is inherited from the Neural Network field. A SVM using a MLP is equivalent to a two-layer, perceptron neural network. So it will be interesting to consider the general performance (accuracy, speed, etc.) between this SVM equipped with this kernel and a ANN model. The two variables κ and c are adjustable, with $\kappa > 0$ and $c < 0$ in general.

The effectiveness of SVM in this work was verified by contrasting the performances/efficiencies of *different* kernels on the dataset. The result obtained here were compared with the one from another machine learning technique: Artificial Neural Network (ANN).

2.1.6 Artificial Neural Network (ANN)

For this work, Artificial Neural Network was used as either the alternative to SVM or a *benchmark* to measure SVM performance. In this chapter, Levenberg-Marquardt optimisation was used as the training function in a Neural Network to classify EEG data. The neural network constructed is a two-layer, feed-forward back-propagation network that has one single output node. Hence the output value regarding to a pattern T is described as [10],[24],[42]:

$$y_1^T = g_O(b_1 + \sum_j W_{1j} \cdot g_H(b_j + \sum_k w_{jk} \cdot x_k^T)), \tag{2.11}$$

- b_1, b_j : the bias
- w_{1j} is the weight of the jth hidden neuron to the single output neuron
- w_{jk} is the weight of kth input neuron to the jth hidden neuron
- x_k^T the kth element of the input pattern T
- g_O transfer function on the output layer - linear transfer function
- g_H transfer function on the hidden layers - sigmoid transfer function

The training performance of the network were evaluate with this error function (mean square error):

$$E = \frac{1}{N} \sum_{k=1}^{N} (y_E - y_P)^2, \tag{2.12}$$

where y_E is the vector of predict outcomes and y_P represents the vector of predicted outcome. The back-propagation training algorithm, being Levenberg-Marquardt optimization, will be represented by the formula [24] :

$$\delta w = (J^T J + I \cdot \mu)^{-1} J^T e \tag{2.13}$$

where J is the Jacobian matrix of the error function calculated in equation(2), μ is the learning rate which is updated after iteration. *diag* being the diagonal of $J^T J$.

2.2 BCI Experiments

2.2.1 Introduction

Human-Computer Interaction research projects rely heavily on the outcomes of the HCI experiments. The fundamental difference between a BCI experiment and other types of HCI experiments is that it is driven by a very *noisy* and *unpredictable* component, the human brain.

BCI experiments were conducted throughout the course of this research. Just like any other HCI experiment, it will involve test subjects to *navigate* through a series of scenarios on a user-interface. As brain patterns being dealt with have yet to be identified, it would be convenient to divide the experiments into two stages: an *Offline* stage and, if possible, an *Online* stage. The main difference separating the two stages is that with an *Offline* system, only EEG signals accordant to various pre-defined scenarios, regardless of the forms/patterns of those signals are captured. That means the *Offline* BCI system will be driven by another mechanism (please see section 2.2.4 for more information) rather than the brain-wave signals. On the other hand, the *Online* system is supposed to be driven solely by the brain-wave signals.

2.2.2 Design Considerations

The experiments involve test participants and a BCI system. The BCI system will present the test subjects with a variety of differentiating tasks and at the same time, record test subject's brain activities with EEG. Among the challenges discussed in Section 2 that one would be facing, the most significant issue is that he will not have the best spatial resolution available with EEG. It would be very difficult to tell if the EEG capture signals accurately reflect the brain activities EEG being tried to measure. Because of that, the BCI systems being considered have to deal with the following issues:

The Noise in EEG Data

BCI systems depend on high levels of *concentration* of the mind to deliver clear enough brain-wave patterns. In contrast, human brains can be very easily *distracted* and *influenced.* Because of that, it is desirable to have a user interface as simple as possible for BCI system. Extra details that draw unnecessary attentions of the mind are *not recommended.*

The Non-stationary Nature of BCI

It is expected to get a large variety of EEG activities even over the *same mental task.* The factors that produce such different signal outputs are very difficult to identify and fully address: types of differentiation, time, human personality, stress level, concentration, etc.

Triggering Mechanism used in Experiment

During the *offline* stage, only the EEG signals was captured for analysing. So in order to navigate around the user-interface, the test subject has to rely on another *trigger mechanism*.

The action of *pressing a key* on a standard keyboard is the preferred option for this purpose. It is known that the brain activities happening during *trivial* action (such as key-presses, where the palms are firmly placed over the keyboard) over the motor/sensory cortex are *less significant* than that performed by *full* hand movements or other alternatives method. The alternative could be participants' pressing a foot pedal to trigger the event. The issue with that is foot movements are not as *natural* and *agile* as the ones performed by hands. Ideally, one would like to be able to *isolate* the EEG activities of the trigger mechanism from the analysis of the captured EEG data. Section 2.6.4 at later stage has a discussion on this matter.

Studied Differentiation Types

The ambition for this research work is to find the pattern of mental differentiating in EEG signals *with or without* regards to the types of the task. However, the difficulties have to be overcome in this work are:

- It is difficult to design BCI experiment scenarios that cover an *extensive range* of such mental tasks.
- The task of categorising and identifying differentiation activities comprehensively is very challenging.
- The "black box" approach being attempted does incur a risk of not being able to fully reason the behaviours of the EEG signals over different task types. Yet, it would be beneficial to build the scenarios around many types of differentiation to gain more background understanding on the nature of them.

Hence the realistic goal set for this research is to focus on only a number of differentiation types. In fact, it is eventually decided to pay special interests in the differentiation that produces *binary outcomes*: *Yes/No* and *Left/Right* are among the possible pair of conclusions. An example for this is to study the EEG activities when a person decides if one item is more than another item in term of a visual attribute. This BCI experiment should only be build around this philosophy.

Using Gaze-tracking Device in Conjunction with EEG

As mentioned above, EEG signals are difficult to work with. If one relied *only* on it could run into the risk of not getting enough data to interpret meaningful information. Having gaze-tracking device that runs during the EEG experiment would add another source of data into this. This is an appropriate proposal given most of the process of differentiating based around visual inputs.
So to be considered a success, the system should be able to cope with those uncertainties and inconsistencies mentioned above. However, it would be near impossible to address all of those factors as in one go. They should be addressed individually as the research progresses.

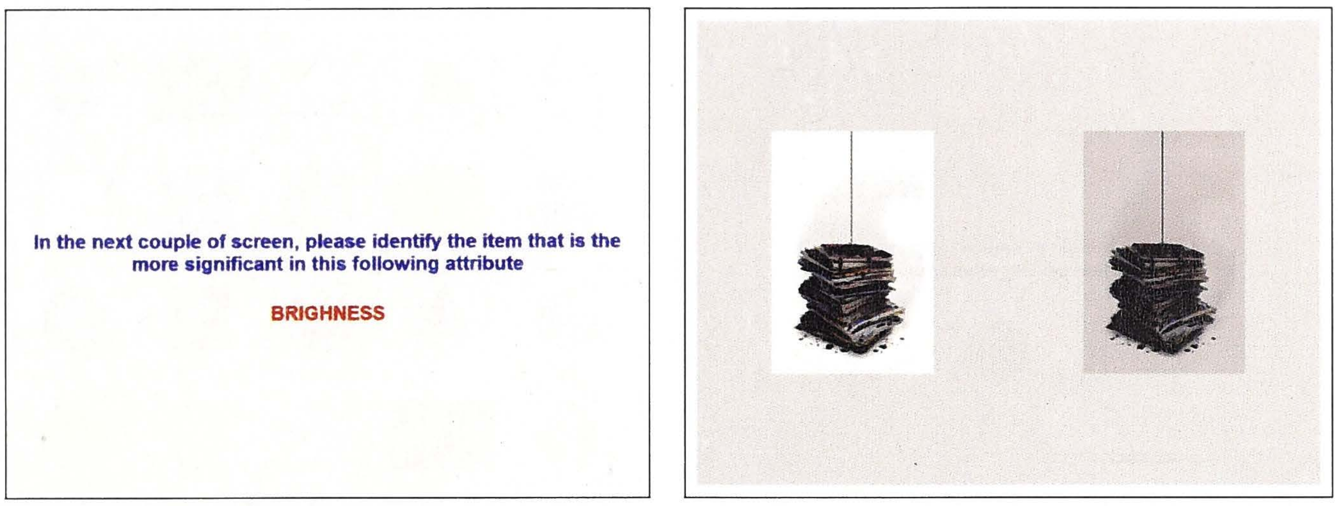
2.2.3 Experiment Details

The BCI experiment involves a subject performing the *three* following tasks to be able to study the *differentiating processes* related to each of them:

Visual Selection Task

This task involves test participants to select images (in his head) based on *visual attributes* that they have been aware of. This task constrains the test subjects into making quick, fast differentiation based on visual clues 2.3(b). The *difficulty* level of this task ranges from easy to medium, depend on how much *different* the two on-screen images are. Below is the suggested list of visual attributes associated to the images that the BCI provides to test subject:

- Brightness
- Colour vs. Greyscale
- Edgy vs. Roundness
- Shades of colours
- Pictures of various types of objects
- Shapes

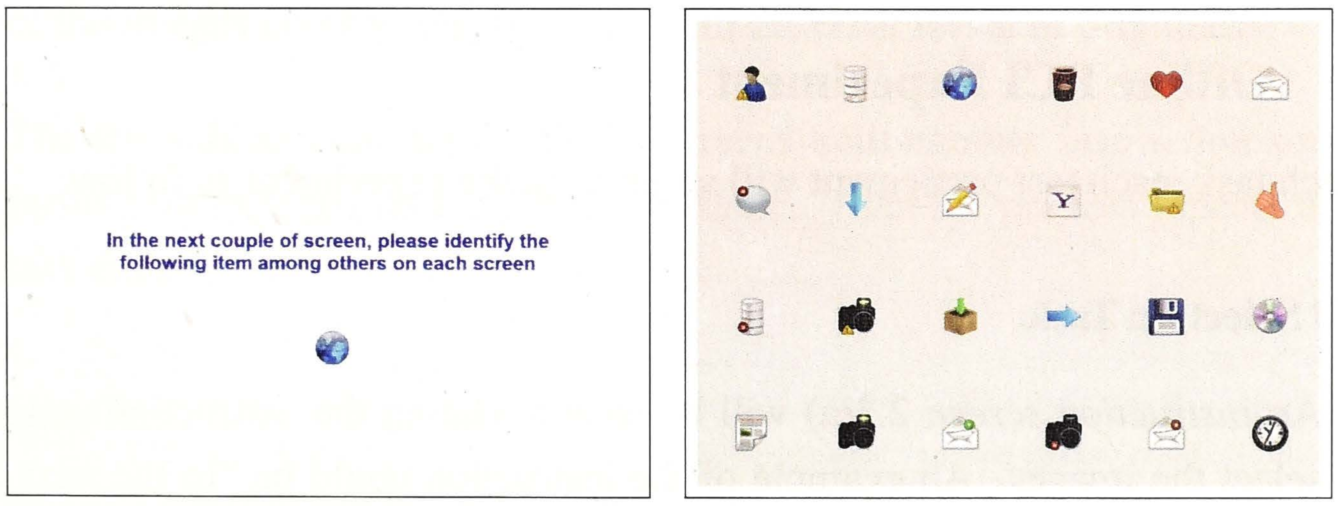


(a) Instruction (b) Typical screen

Figure 2.3: Screenshots of Visual Selection task

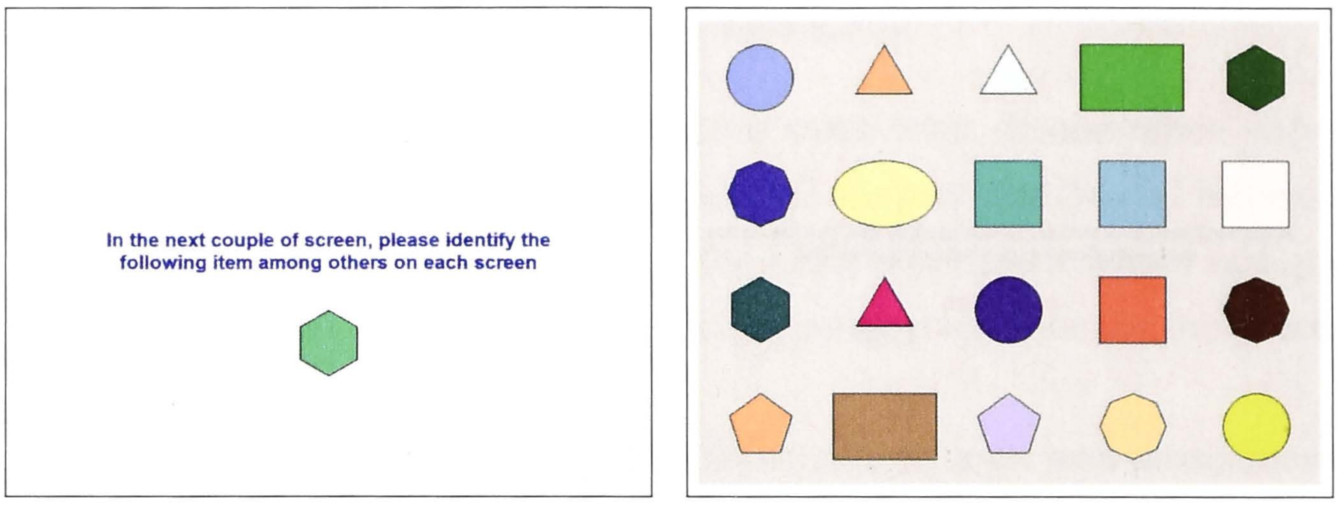
Searching Task

This task involves test participant to identify and select *specific* images (in their heads) that are placed among other images on a series of screens. For each case, the test participant is informed about the *targeted* image before he could be able to find it on each screen 2.4. This task *guides* the test subjects into performing the combination of identifying, comparing and finally differentiating whether an image is the *target* image. The level of this task ranges from easy to medium, depends on the number of items on the screen. Another flavour of this task will also be part of the experiment. However, instead of having *images*, a combination of geometric shapes with different colours as the items to be investigated on screen 2.5 was used.



(a) Instruction (b) Typical screen

Figure 2.4: Screenshots of Searching task (With Pictures)



(a) Instruction (b) Typical screen

Figure 2.5: Screenshots of Searching task (With Shapes)

Advanced Searching Task

This task involves test participants to find and select images (in their heads) that are placed among other images on a series of screens. However, unlike the above *searching task*, the test participant *does not know* about the *targeted* images. The only condition guaranteed here is that for each case, the*targeted* image *always* appears on every screen of the series 2.6(b). So they have to *find* it by memorising and identifying a particular image that always appears on each screen. This task *guides* the test subjects into performing the combination of following actions: identifying, memorising, comparing, eliminating (deducing) and finally differentiating on whether an image is the *targeted* image. At the end of the series, the test subjects will be asked to choose the image that they thinks is the *target* out of the four given choices 2.6(c). The level of this task ranges from medium to hard, again depends on the number of items on the screen.

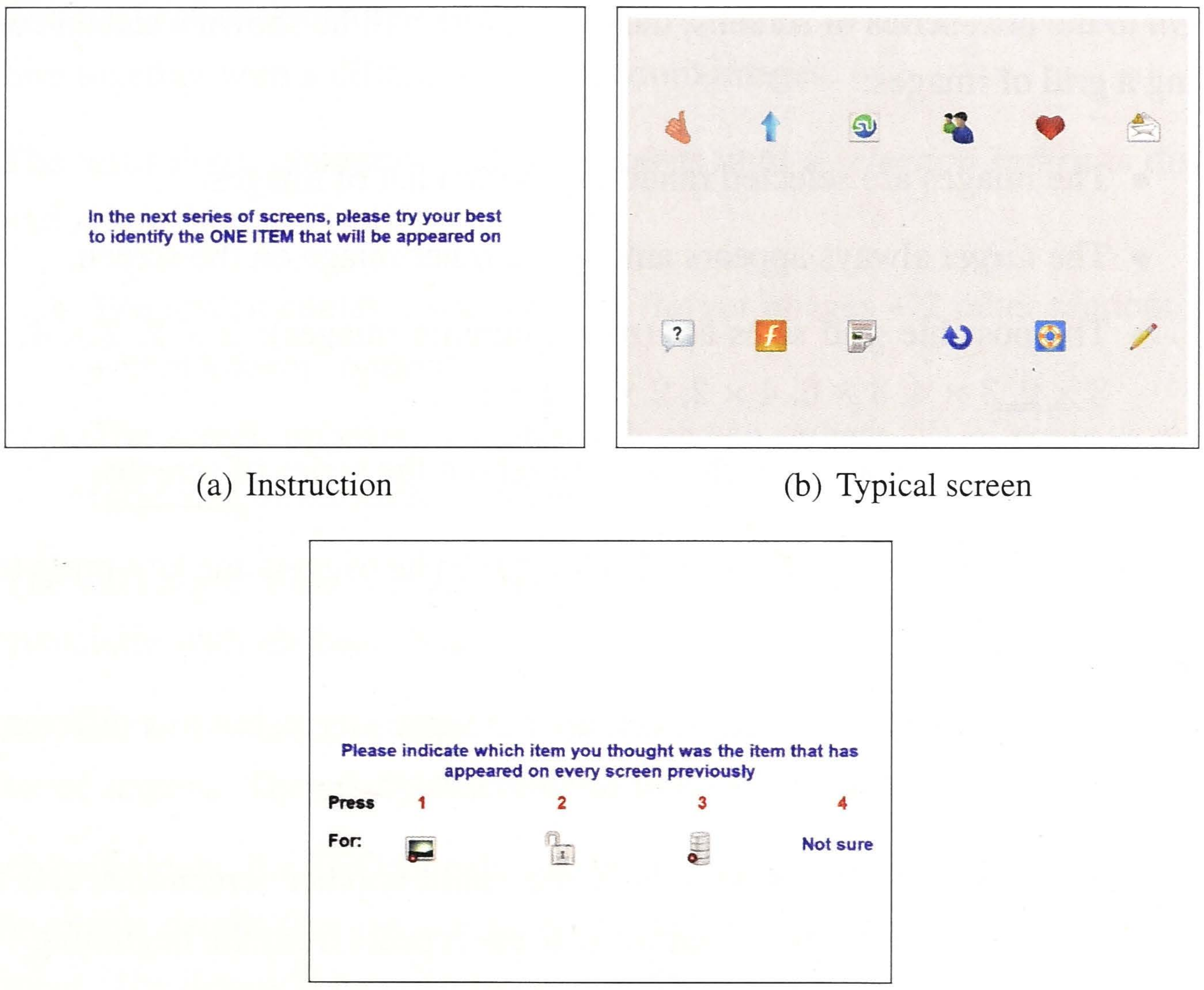
2.2.4 Offline BCI Experiment

For each task, each test participant will sit through the experiment as follow:

Visual Selection Task

1. An *instruction screen* 2.3(a) will be shown, stating the instructions on how to select the images. An example of the instruction could be "In the next couple of screens, please identify the item that is the most significant in this following attribute: brightness". This screen also instructs clearly that the instance the test subject has made his choice, he has to press the *spacebar* key to indicate so.

2. On the next screen, the test subject will be shown a screen 2.3(b) comprising of



(a) Instruction

(b) Typical screen

(c) Question screen

Figure 2.6: Screenshots of Advanced Searching tasks

two images. To continue with the example, one of the images is brighter than the other. He chooses in his head then he presses the key to go to the next screen.

3. The set of screens that follow will still be about the difference in the brightness of the images given to test subject - with different levels of brightness.

4. The test subject continues with the screens until another *instruction screen* informs him to focus on a *different attribute*. Then the process repeats until the task ends.

Searching Task

1. The test subject will be shown a screen having a *target* image together with the instruction. An example of that could be "In the next couple of screens, please identify the following item among others on each screen <image>". This screen also instructs clearly that the instance the test subject has made his choice, he has to press the *spacebar* key to indicate so.

2. On to the next series of screens, the test subject will be shown a screen comprising a grid of images:

    - The images are selected randomly from a list of images.
    - The *target* always appears among the other image on the screen.
    - The possible grid sizes are (row x column images): 2×2, 2×4, 2×6, 3×2, 3×4, 3×6, 4×2, 2×4, 4×6.
    - The grid size does not change throughout the series of screens.

3. The participant finds the *target* in his head then he triggers the key-press to go to the next screen.

4. The next set of screens will still contain the same *target*, but in a different location together with a different set of random images.

5. The test subject continues with the screen until another screen instructs him to focus on a different *target*. Then the process repeats from the beginning.

Advanced Searching Task

1. The test subject will be shown a screen showing the the instruction. An example of that could be "In the next series of screens, please try your best to identify the ITEM that will appear on every screen.". The test subject *does not know* what will that item (or *target*) be. This screen also instructs clearly that the instance the test subject has decided to move on to next screen, he has to press the *spacebar* key to indicate so.

2. On to the next series of screens, the test subject will be shown a screen comprising a grid of images.

    - The images are selected randomly from a list of images.
    - The *target* always appears among the other image on the screen.
    - The possible grid sizes are (row x column images): 2×2, 2×4, 2×6, 3×2, 3×4, 3×6, 4×2, 2×4, 4×6.
    - The grid size does not change throughout the series of screen.

3. The participant identifies the *target* in his head (by memorising and deducing the possibilities), then he triggers the key-press to go to the next screen.

4. The next set of screens will still contain the same *target*, but in a different location together with a different set of random images.

5. The test subject continues with the screens until a *selection screen* is displayed and asking him to identify the *target*:

   - The screen contains four choices (*target* images + 2 other random images + "not known" option).
   - The screen appears after the Nth screen, where $N = rows + columns$ (grid size of the screen series).

6. The test subject then indicates his choice by pressing either key 1, 2, 3 or 4 (that associates with his best choice).

7. Another screen instructs him to refocus on finding another *target* with a different set of screens. Then the process repeats from the beginning.

For each participant, the EEG signals will be recorded throughout the session together with the events details (timing, screen type, etc.). The data will be analysed after the experiment. The data will be analysed to identify the best method of classifying them. Ultimately, these methods would be used in the online system. The detailed plan for analysing data will be discussed later on. The synchronisation of events happened between the components are handled using software with the timing difference (offset) are took into account.

2.2.5 Hardware Option 1 - BioSemi ActiveTwo

The ActiveTwo system is a multi channel, high resolution *biopotential measurement* system for *research* applications provided by BioSemi [3]. The system is a development from ActiveOne system, the first commercially available system with *active electrodes*[3]. Please refer to section **??** for more details on the device. Here are some of the information of this device that is relevant to the project:

Accuracy and Performance This device belongs to the class of medical-graded devices. With a very high and customisable hardware capture rate (up to 16 kHz/channel), this device is expected to be the one that delivers most precise and accurate signals. The use of *fiber-optic* for signal transmission to *ADC* box has signified the effort of the manufacturer in delivering the best signal-to-noise ratio among competitive devices. However, access was limitted to *16 channels*, a subset of its full capability of 256 channels. It is clear that such potent device was not utilised at its full potential.

Usability The device requires careful test participant preparation. Largely due to its operational nature, there are some inconveniences in using this device. The use of conductive gel and the the availability of only one head cap (which has to be shared among test participants) have limited the turnaround of experimental trials that can be completed in one day.

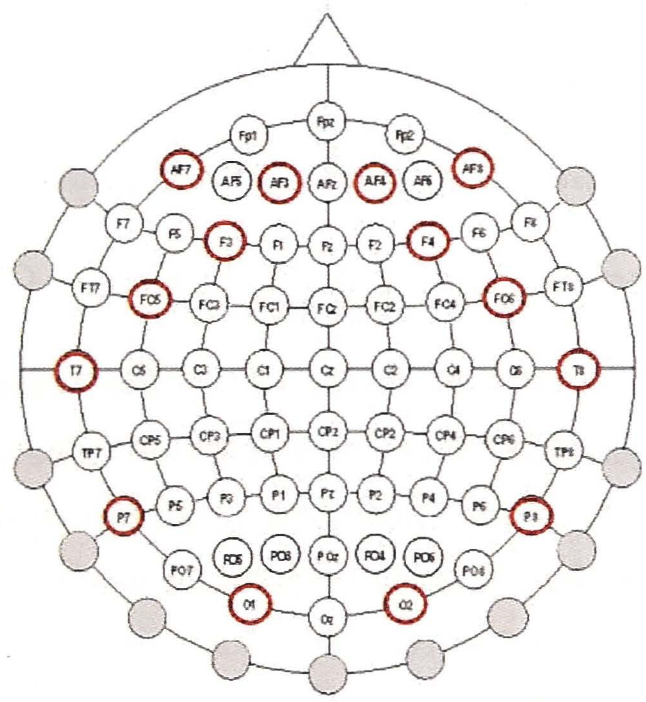
Interface and software options The device utilises an application called *ActiView* as the acquisition software. It can only be customised by the LabVIEW graphical development environment. Unfortunately, access to the LabVIEW suite was not available and ActiView was found not suitable for this research needs. The only alternative to ActiView is to write specialised acquisition program with very little support from BioSemi. Please refer to section 2.2.8 for more details on other issues with the acquisition software aspect of this device.
Since the start of the project, ActiveTwo was always the device of choice for EEG measurement. However, from 2011 onward, the use of ActiveTwo as the EEG measuring device has been postponed. Emotiv's EPOC headset has been used as a replacement instead.

2.2.6 Hardware Option 2 - Emotiv EPOC headset

This EEG headset is essentially the EEG solution provided by Emotiv [4]. Unlike BioSemi device, the ADC component is built into the headset hence the need for fiber-optic cable is no longer required (see 2.2.5). According to Campell [11], the headset transmits encrypted data wirelessly to a Windows based machine; the wireless chip is proprietary and operates in the same frequency as 802.11 (2.4Ghz) standard. Figure 2.7(a) shows the headset with its electrodes. The 14 electrodes are placed roughly according to international 10-20 system. Consequently, they are labelled as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4. Figure 2.7(b) show the location of the electrodes with red highlighted border.

Accuracy and Performance Unlike the offering from BioSemi, EPOC can only capture at 128 samples/second. There is no direct comparison of the capture quality between EPOC and ActiveTwo. However, the trials with the device within this facilities has shown that the device is more than capable in term of Signal-to-Noise ratio (with or without filtering). The only concern is the low sampling rate: with the *differentiating processes* are assumed to be short, the low sampling rate means that one may not be able to capture the activity as accurately as he would like.

(a) Emotiv EPOC headset



(b) Electrode locations according to 10-20 system

Usability The device still requires careful test participant preparation. The electrodes, which are felt-based, are placed directly on the scalp. They *do not* require the use of *electric conductive gels* as in the case of Biosemi's ActiveTwo. This is an advantage of this device over BioSemi's offering. The efforts in preparing the test participants can be significantly reduced because of that.

Interface and software options Acquisition software comes with the device is *Emotiv TestBench*. Its functionalities are almost identical to ActiView of BioSemi. Just like the other acquisition suite, it also handles event triggers via serial port and can write EEG signal to EDF files.

From 2011 onward, Emotiv EPOC has been used as the EEG measuring device. This decision was made mostly for convenience and other practical reasons: efforts on setting up equipment, driver supports and trial turnaround time, etc.

2.2.7 Software Option 1 - BCI2000 Project

Introduction

BCI2000 [5] is a framework for BCI systems. Its modular deign is flexible enough to allow researchers to customise different BCI systems that suit their needs. In general, BCI2000 provides these modules that form a complete BCI system:

1. Source Module (Capture source / Storage)

2. Signal Processing Module (For example:band-pass filters, frequency transformations, etc.)

3. Classification module (For example: linear classifier, pattern recognition, etc.)

4. Application Module (For example: GUI, visual stimulus, audio stimulus, etc.)

Each of the section above can be customised, taken out or replaced to suit specific needs. This framework was originally chosen because that would save time and effort of developing one whole BCI system from scratch. Another beneficial factor about this framework is that it comes with a set of useful and well-written software modules that can be re-used.

Strategies

Here are the list of items that are going to be reused from the current BCI2000:

1. The core component (controller).

2. Band-pass filters.

3. The capture module for BioSemi2 device.

4. BCI2000 file writer.

Here are the list of items that are going to be modified or replaced with in BCI2000:

1. GUI interface: Needed to be modified to suite the purpose of my experiment.

2. Signal processing: Will use an Fast Fourier Transformation module in place the default options.

3. Linear classifier: Will supplement a Support Vector Machine module in place of the default linear classifier.

Issues with BCI2000

During the course of customising the BCI2000 framework, quite a few issues had became apparent:

- **Inflexibilities to keep up with the requirements**: The aim is to produce three small set of experiments described in 2.2.3. Those three tasks required customising quite a few changes in the *GUI* and the *Input* components. However, as these customisation tasks progressed, it is discovered that BCI2000 was not as flexible as original anticipation. It resulted in a significant amount of effort in customizing the BCI2000 framework that far out-weighed the benefits the framework offered.

- Device drivers for BioSemi2: The correctness of current implementations for the BioSemi device's capturing module is still an uncertainty. Furthermore, the documentation and support for this work are inefficient. As from 2011, Biosemi has updated the driver of the BioSemi2 device, rendering the current capturing module of BCI2000 *non-functional*. A major update effort for the module is required. Currently, no active effort has been allocated on this work from the BCI2000 community.

- Device drivers for Emotiv: The use of Emotiv device in this research work requires an implementation of capturing module for the Emotiv device. The status of the work on this is unclear.

Out of the above problems, the inflexibility of the BCI2000 suite is the major factor. The long development cycles and high dependencies in codes have made other alternatives being take into consideration.

2.2.8 Software Option 2 - Custom Built Application

The alternative to the solution with BCI2000 is to develop a lighter-weight solution. Along side with BCI2000 work, a light-weight tool that can perform similar functionality has been developed.

- This software is developed with C# and .NET platform, with Visual Studio being the development environment.

- For the *Offline experiment*, this software does not capture EEG signals. Instead, the *acquiring application* provided by device manufacturers perform that responsibility. The software, however, communicates with the acquiring component by *events*.

- For the features that are still lacking from the BCI2000 framework (filters, classifier), there are options with third-party solutions. .NET application can interact with these components via .NET C/C++ interface.

Without most of the unnecessary modules and code dependencies, the development time has been significantly reduced. Many restrictions in designing GUI components and experiment scenarios are also removed, resulted in much more flexibility in tailoring the experiment according to requirements.

From 2011 onward, this collection of tools/modules has replaced BCI2000, at least for the Offline phase described in 2.2.4 to reduce the efforts to get to the initial analysis/validation milestones. With better consideration and preparation, the Online phase could still benefit from the BCI2000 solution.

2.3 Signal processing

Signal processing, in this context refers to the signal *pre-processing block* for BCI system. EEG signals, by nature, are considered both very noisy and unpredictable. Filtering out unnecessary bits of data as well as identifying principle *feature set* are the *de facto procedures* when operating EEG signals.

Filters Band-pass filter was used to remove the irrelevant frequency ranges. The frequency range in 0Hz - 100Hz is considered adequate which is a little bit larger than the norm but it should keep enough useful information to be utilised.

Frequency series transformation (FFT) The EEG signals, having being filtered, will then be transformed into the frequency series. This is the final stage of reprocessing. I chose the FFT method for this purpose and this will be used on both the online and offline system. The output of this transformation will be the representation of the EEG signals in terms of *activations* (altitude) in the frequency bands. These values are expected to be used as data for machine learning training/classifying.

2.3.1 Fast Fourier Transform (FFT)

Studies of EEG are generally driven by the states of five different frequency bands: *delta*, *theta*, *alpha*, *beta* and *gamma*. Given the EEG signals are not represented in that form by default, it is just natural to transform the raw EEG signals into the frequency domain using FFT. The outputs of FFT will be used to identify the EEG activity in each of the above bands. Refer to section **??** for more information on FFT. FFTW3 library was used for this purpose. For that, the library was configured to operate in *Halfcomplex-format FT* mode. This mode receives real input values and converts those into half-complex value frequency-domain values. About the FFTW3 library: [20] FFTW does not use a fixed algorithm for computing the transform, but instead it adapts the DFT algorithm to details of the underlying hardware in order to maximise performance. For this case, FFTW provides fast planners based on heuristics or on previously computed plans.

2.3.2 Key-press or Not Key-press

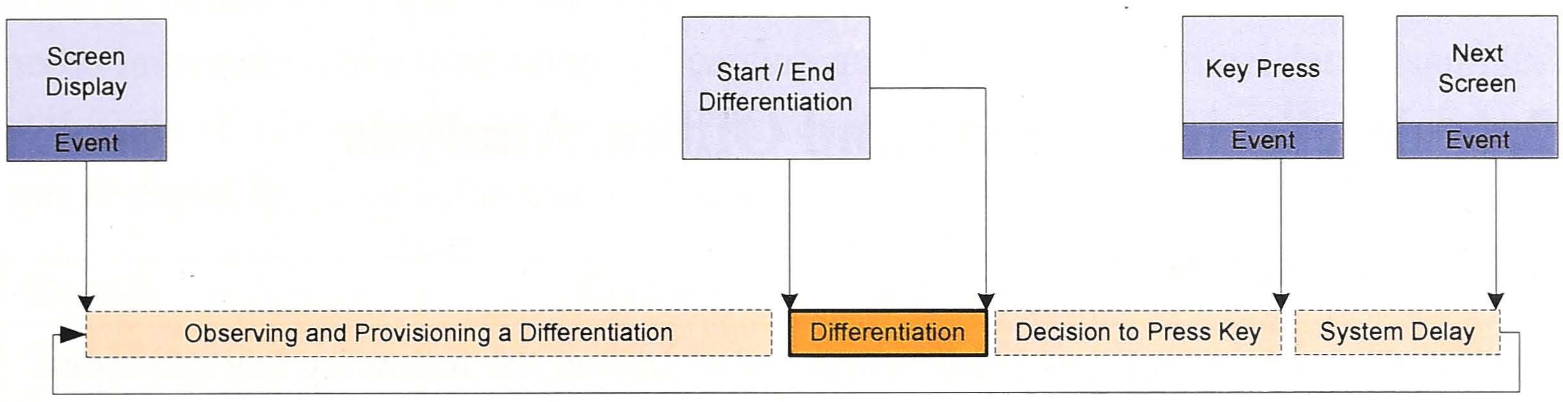


Figure 2.7: Event sequence of a typical experiment unit

As one can see from Figure 2.7, the period of *differentiation* (yellow box with bold border) is the period that one cannot pin down with event markers properly. This is a design *shortcoming* that could not be properly addressed. A test subject, according to given instructions, once he finishes the mental task should press the spacebar key *as quickly as possible*. However, how *quick* is *quick enough*? And how can one to correctly measure that? In other words, how could he separate the *differentiating process* and the *key press activities* without the help of event markers on the EEG signals? There are a few solutions that are worth considering. The common goal for these solutions is to effectively detect the brainwave activities of the *key-press* events:

- **Study of sensorimotor evoked potential** Key press involves activating the sensorimotor rhythm to *some degree* from the sensorimotor cortex. The frequency of the rhythm is between of 8 and 12 Hz and is considered very close to the mu rhythm [8]. In theory, one could study the signal patterns around the *KeyPress* event in that frequency range to confirm the *existence* of that rhythm, then he can *eliminate* it out from the analysis. However, the correctness of that approach, together with the solution chance of success is quite *uncertain*.

- **Study of key-press activities independently** Another solution is to study the brainwaves of key-press activities *outside* the experiment restrictions, i.e. to study the EEG signal of test subjects performing key-presses *without* the constraint of having to make a differentiation prior to it. These brain-wave patterns will be studied, compared with the data captured during the experiment and the removed from the final analysis. Having said that, the efforts in realising this solution are quite significant.

Individually, both of the solutions above is a substantial unit of work and should be carried out with careful considerations. There is also a possibility that neither of the techniques would help improve the situation.

2.4 Preliminary Trial and Offline Analysis

2.4.1 Background

Data analysis of the test data would allow me to answer the following key questions:

- If there is a pattern of the EEG signals that exists across different types of differentiation at the very moment it is made.
- If there is a pattern of the EEG signals that exists across a specific type of differentiation at the very moment it is made.
- If there is a pattern of the EEG signals that exists across a test subject/type of test subject at the very moment it is made.
- Are there any interesting and distinctive features that one could benefit from?

The EEG signals, being recorded by the electrodes placed on the human scalp are, in fact, the summation of all the electric activities happen within the brain. The EEG equipment in use (Emotiv EPOC) has 14 EEG channels - and it is important to identify which of those channels carry the information that are interested(*featured* channels). For that purpose, the use of spectral analysis on the signals around the *event markers* was proposed. Ultimately, only the feature channels were used in data classification process.

At this stage, there is also an opportunity to experiment with machine learning algorithms (see section 3.6.2 and 2.1.5 on the EEG data just to have some *indication* on the effectiveness against classifying the processed EEG signals.

2.4.2 Pre-processing

Once the preliminary trials have concluded, the EEG data is recorded and stored in EDF files. Originally, the intension is to run the *offline* experiment on all *six* test participants and performed the analysis work as described in this section. Out of the six participants took part in the trial, three sets of data have been considered severely corrupted by noise. Hence they were deemed not appropriate to perform studies on. The remaining three datasets are considered suitable for further investigations.

For each person, there is one EDF file for his recording. Even though each file include a total of 35 data channels, there are only 14 channels that are relevant to the analysis (Channels 3 to 16). Besides those 14 channels, the others contain miscellaneous information like timestamp and capture quality, etc. The channel data is sampled at the rate of 128 samples per second. As for the EDF specification, the last channel was reserved for *event* information. The *events* are coded as described in Table 2.1.

Events	**Codes**	**Events**	**Codes**
Test Start	0	Task Start $_{(1)}$	$30 + taskcode$
Instruction Display	1	Instruction Hide	2
Next Screen	3	Screen Display	4
Task End	5	Test End	6
Question Show$_{(2)}$	7	Question Hide$_{(2)}$	8

Table 2.1: System generated Events and their equivalent codes used in EDF files
(1). Each task type has an type number (for instance 1 for Task 1 and so on). Resulted in the event codes are 31, 32, 33 etc. (2). Only applied to Task 3, where there will be questions given to the test subjects.

2.5 Preliminary Trials: Initial Observations

Figure 2.8 shows the plots of Event Related Potential activities within the *epochs* created around the *Next Screen* event. Please note that this event is raised whenever the test participant performs the key-press action to indicate he has made a differentiation. This *initial* observation disregards the concerns mentioned in section 2.3.2 for the time being.

Within Figure 2.8, the *key-pressed event* occurs at offset $0ms$. There are a few things that can be observed from the figure:

- The scalp map(s) of each epoch figure consistently indicate that the majority of activations (spikes) of the EEG activities happen around the *frontal lobe* section of the brain, as depicted by figure 2.9(a).

- The processed signals of these activities are depicted in Figure 2.9(b). From the figure, the EEG activities of the three frontal channels *AF3, AF4, F8* can be easily identified as the reason behind these spikes.

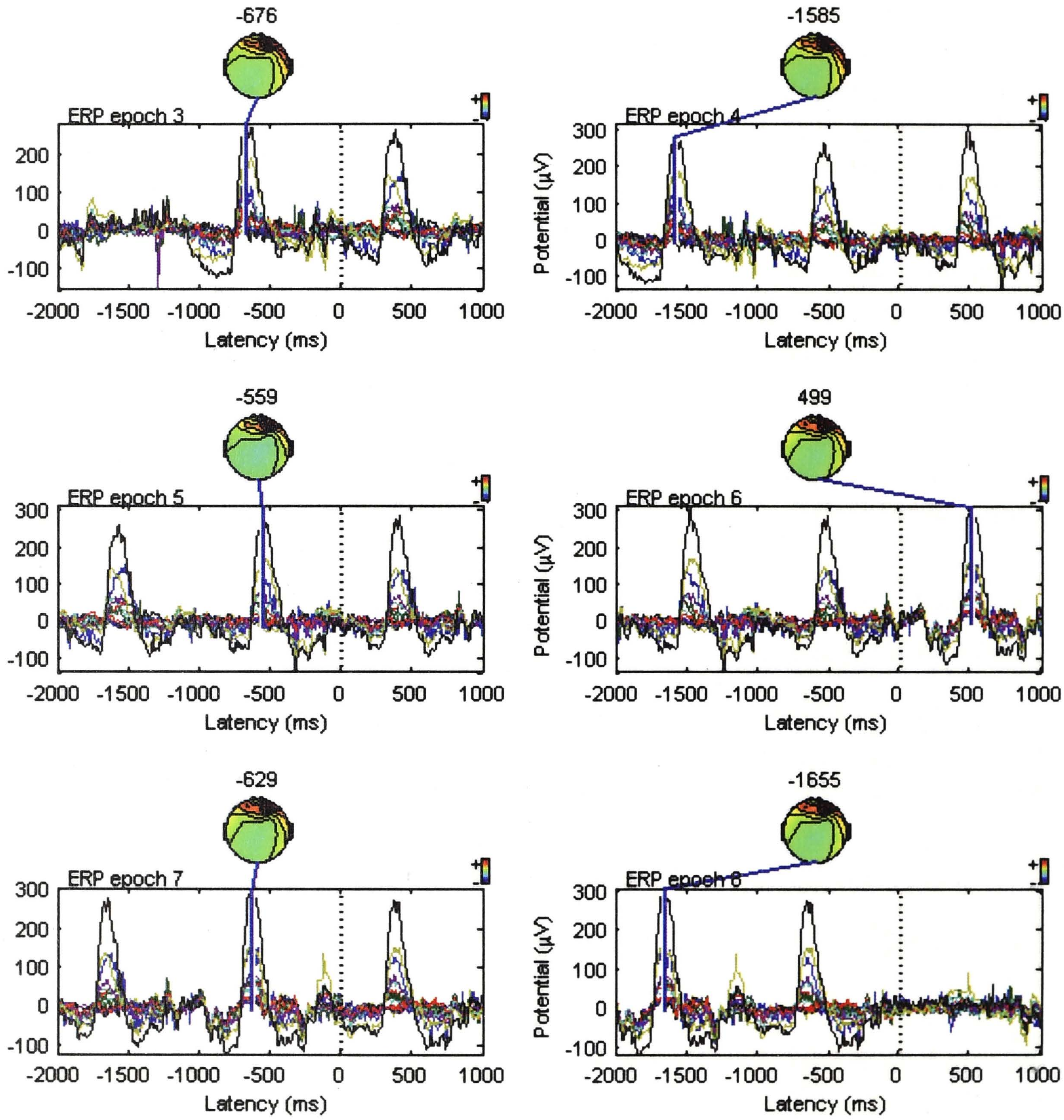
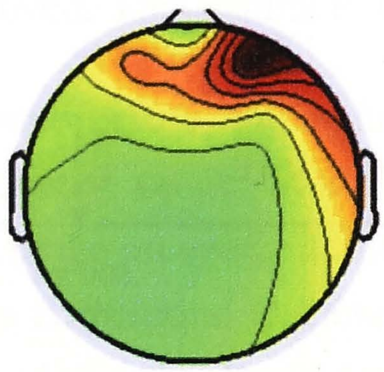


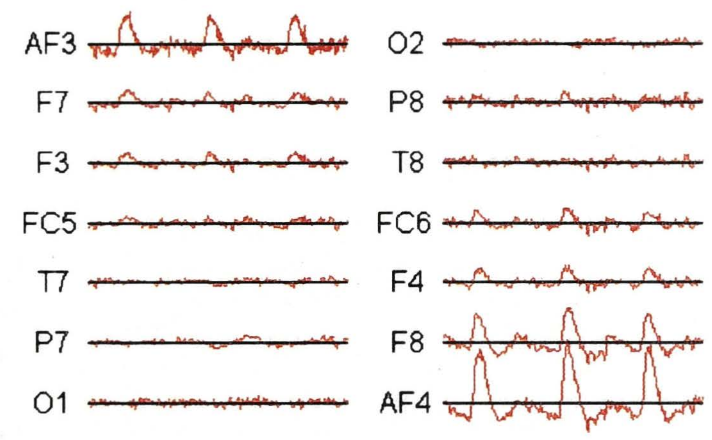
Figure 2.8: Plotting of ERP data with scalp maps for Epochs

- Signals spike at around latency *-500 ms*: happened in *all* epochs 3, 4, 5, 6, 7, 8.
- Signals spike at around latency *-1500 ms*: happened in epochs 4, 5, 6, 7, 8.
- Signals spike at around latency *500 ms*: happened in epochs 3, 4, 5, 6, 7.

Because anything happens after the *event* (post *0 ms* mark) 2.8 is considered irrelevant to the *differentiating* process, that leaves the option to only statistically study the EEG activities around the *-0 ms* and *-1500* marks. As the event at 0ms can only happen after



(a) EEG activations heatmap of one epoch



(b) EEG activities depicted in the scenario in Figure 2.8

Figure 2.9: Brain activities of a differentiating task: Activations and Raw signals.

the differentiating process has concluded, this leads to the following hypothesis being proposed:

Hypothesis 2.1. That certain spikes in EEG signals have some connections with the *differentiating process* that happened within the same time frame. If one to perform a statistical analysis on the signals in between the time range when those spikes tend to occur (between *0 ms* and *-1500 ms* marks of the aforementioned *even markers*), he could find out if he can identify the *differentiating process* from EEG signals in those time frame.

2.5.1 Statistical Analysis

Signal Preparations

The work in this section is done with the help of EEGLAB [2], an interactive Matlab toolbox for processing continuous and event-related EEG signals. In order to be further analysed, raw EEG signals are pre-processed as followed:

- Select the 14 channels from the EDF file.
- Load the location coordinates (in 3-D Cartesian coordinate) for all channels.
- A band pass filter was applied on the EEG signals which admits only the frequencies from 3-30 Hz. This is done to reduce noise at uninteresting frequency ranges.
- In this analysis, onlyone type of event was considered, the *Next Screen* 2.1.

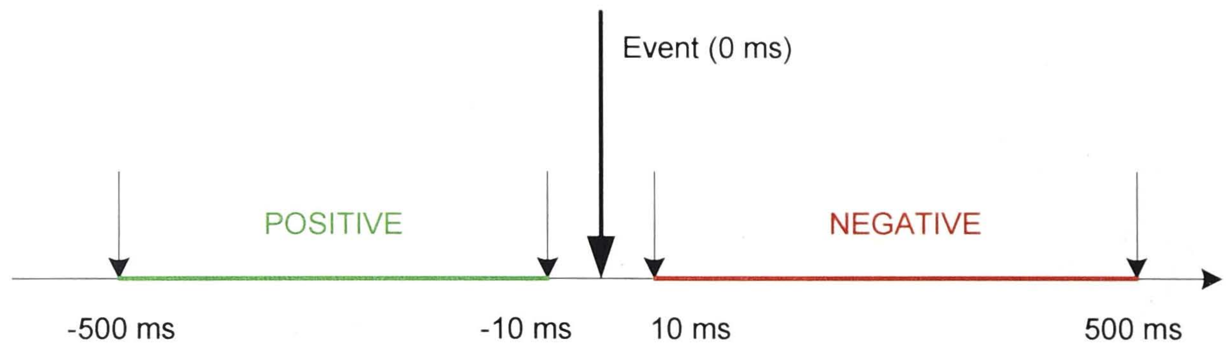


Figure 2.10: Relative positions of Positive and Negative epochs

- Continuous EEG data is extracted into *two* groups of *short epochs*. One of them consists of epochs that happen *before* the event (positive epochs) while the other involves epochs that happen *after* the same events (negative epochs).

- Remove mean baseline values of data epochs(base on the difference in values created by low frequency drifts or artifacts compared to values around each epochs)

The task of extracting continuous EEG data into epochs is explained in more detail here. The reason for using only one type of event is to *minimise* the number of experimental variables as well as to *simplify* the analysis process. As mentioned before, there are two groups of EEG epochs created after this tasks.

- **Positive epochs** Represents the epochs that contain the *differentiating process* signals. The time range offsets for each of these epochs is [-0.5 -0.01] seconds i.e. range from 0.5 seconds till 0.01 seconds right *before* the event.

- **Negative epochs** Represents the epochs that do not contain the *differentiating process* signals. The time range offsets for each of these epochs is [0.01 0.5] seconds i.e. range from 0.01 seconds till 0.5 seconds right *after* the event.

Figure 2.10 also describes the positions of each epoch types in relation to the event's position.

The final stage of processing the EEG data is to transform the EEG signals from *time domain* into *frequency domain*. The details are as follows:

- Fast Fourier Transformation is used on the three channels: *AF3, AF4, F7* and *F8*. This channel selection is based on the discussion in the previous section 2.5.

- Each epoch is about 0.5 seconds long, which result in around 64 points at 128 sampling rate. Hence, the FFT window size is 64 (the next higher power of 2 to the amount of data points in each epoch).

- Because FFT transformed data is symmetrical, the second half of the data was removed. This results in FFT *vectors* of 32 length in data points.

As mentioned before, the signals are band-passed to retain the signals within the 3-30 Hz frequency range. Studied have found that most of brain rhythms occur in that frequency range on *awake* adults [48]. Since that is the case, only the *alpha*, *theta* and *beta* waves correspond to those awake brain activity are to be studied closely for the remaining of this chapter.

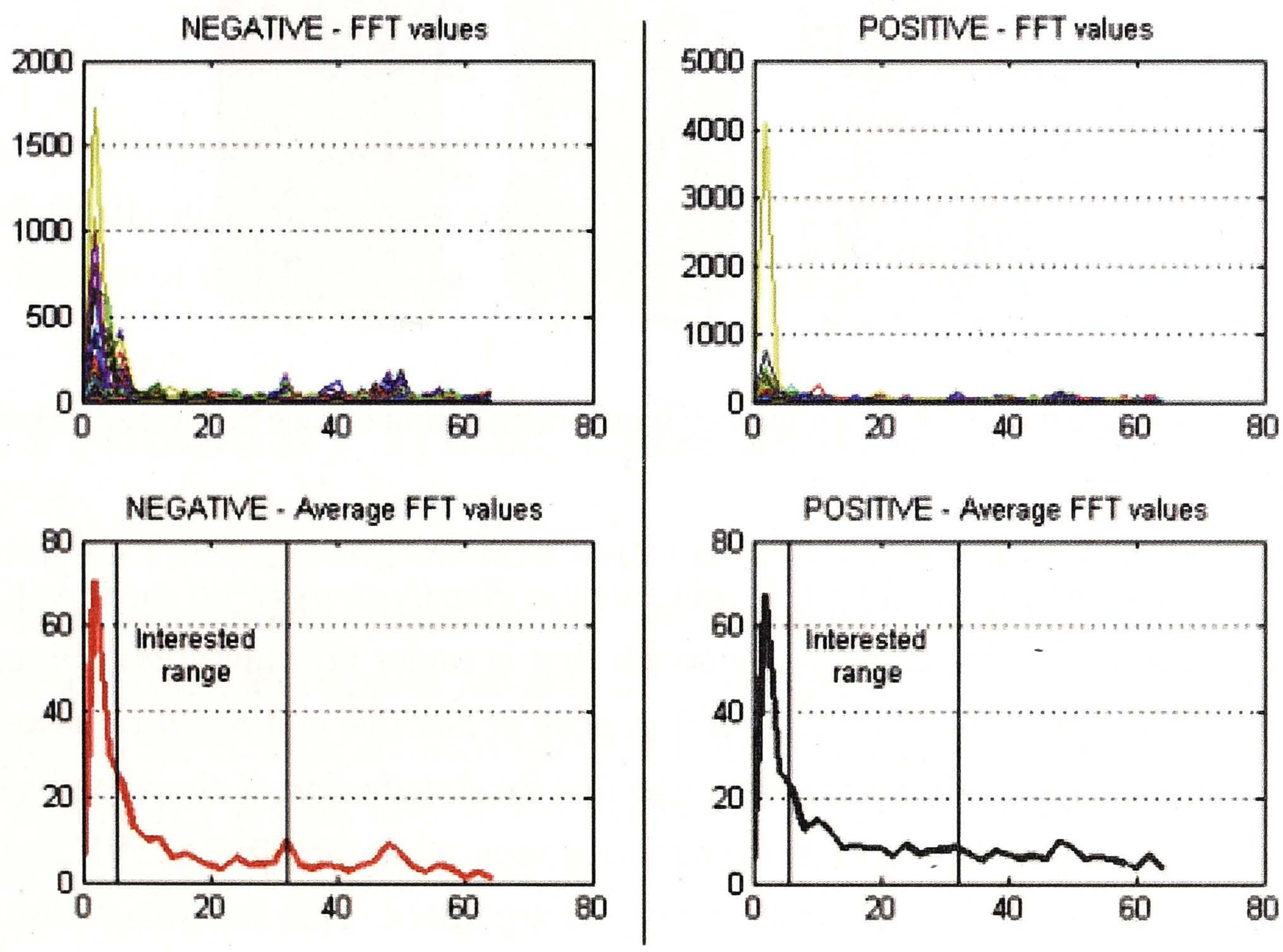


Figure 2.11: Comparison of EEG frequency activations: Positive vs. Negative

Positive vs. Negative Cases

Figure 2.11 demonstrates the frequency transformation values of individual *epochs* as well as the *averaged* values of all epochs. The epochs were extracted from the recording of a participant performing the *searching* task (section 2.2.3). The plots have shown that, visually, there are *subtle* differences in frequency activations between the Positive and Negative cases. It is debatable that those differences are significant enough, however. To further analyse this data, the data points were grouped into three bands (Alpha, Beta and Theta) based on their equivalent frequency (Hz). For each band, the *mean* value of all data points within it was calculated. This is done for every epochs created so far. As before, they are then being split into two groups:Positive and Negative. Figure 2.12 contrasts the differences between the *positve* and *negative* epochs

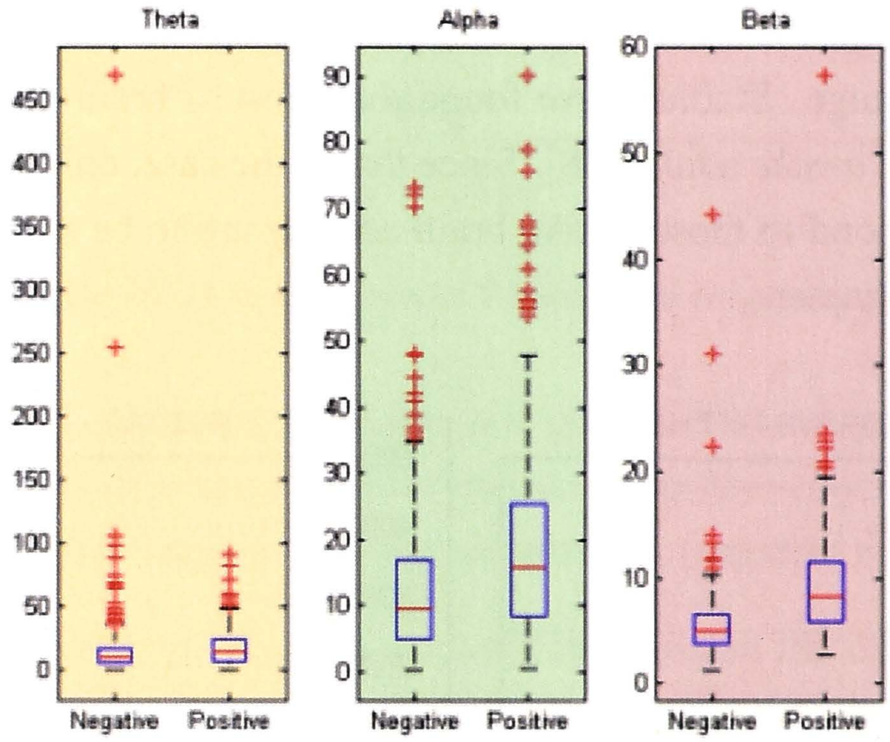


Figure 2.12: Comparison of EEG frequency bands distribution on one participant: Positive vs. Negative

in signal activations of each band. It presents the *distributions* of *all* mean values calculated for each EEG band for *every* epoch that is under consideration. Figure 2.12 shows that there are some distinctive differences in *positive* and *negative* cases across three EEG bands. One can tentatively argue that the*distribution* of signal activations in the three bands (alpha, beta and theta) can help separating the NEGATIVE and POSITIVE cases. Table 2.2 further consolidates that argument. This *statistical observation* so far has set some ground for further analysis. One would still like to know if those discrepancies he has identified are *significant* and *consistent* enough across the dataset: subject to subject and tasks to task.

2.5.2 Discussion

The initial statistical analysis above has signified a few things. Firstly, there are subtle differences in EEG activities within the to the two types of epochs (Positive vs. Negative). Whether these differences are *significant* enough to be the deciding factors in identifying the *differentiating* patterns is yet to be confirmed. Beyond the scope of this analysis, there are more advanced and capable methods to study the statistical nature of the EEG data. However, the use of *statistical machine learning* techniques to confirm the statistical significance of the aforementioned differences was the chosen option. If *machine learning classifiers* could achieve a high enough accuracy, cross-validated,

	Subject $1_{(*)}$	Subject $2_{(*)}$	Subject $3_{(*)}$	
Task 1	9.47 ± 4.37	8.04 ± 5.52	8.56 ± 15.36	Positive
	6.22 ± 2.79	4.57 ± 4.25	5.95 ± 16.12	Negative
Task 2 (Shapes)	7.74 ± 4.44	7.85 ± 6.61	19.49 ± 30.46	Positive
	4.63 ± 3.15	4.68 ± 6.31	15.65 ± 24.69	Negative
Task 2 (Pictures)	9.35 ± 5.33	4.92 ± 2.93	16.08 ± 41.89	Positive
	5.68 ± 3.80	1.67 ± 0.86	10.33 ± 23.12	Negative

Table 2.2: Measurements of activations within *Beta* band frequency.
(*) $mean \pm standard\ deviation$

with the same dataset, one would be confident enough to confirm this *hypothesis*: "One can statistically identify the *differentiating process* from EEG". That leads to the *classifying* section of the chapter.

2.6 Preliminary Trials: Classification

In section 2.5.1, the use of *statistical machine learning techniques* as the alternative for the traditional analysis described has been mentioned. This is a good opportunity to try these techniques because the ambition is to be able to utilise *machine learning techniques* as the *classification* component in BCI system building blocks. The two stages of this BCI system involve the *machine learning* techniques in the following manners:

Offline stage Data is mainly used for training and verifying the performance of the classifier. One will try to obtain optimal results for each differentiating types - even if that means he would end up having different sets of training outcomes.

Online stage Data being captured in real-time will be classified using the parameters obtained from the training stage. Hence that completes the online BCI system. *Support Vector Machine* and *Artificial Neural Network* are the tools that the focus is on.

2.6.1 Methodology

Similarly to the previous analysis, it begins by extracting EEG data into epochs (as described in 2.10). Hence for the purpose of *labelling* the dataset:

- **ONE** for epochs that are within the POSITIVE durations.

- **ZERO** for epochs that are within the NEGATIVE durations.

Then, for each epoch, the raw EEG signals from *four* EEG channels *AF3, AF4, F7* and*F8* are pre-processed identically to the process described in section 2.5.1, up to the step where the FFT signals are placed into EEG bands. For each channel above, the *mean* power values for each EEG frequency band was computed. The selected values wrere calculated from the three bands *Alpha, Beta* and *Theta* as part of the feature vectors. This results in 12 features for each sample (4 channels $\times$ 3 mean values).

Normalisation of Vectors in Input Space

Before being trained with SVM, each value of the *Input-spaced* vectors will be *normalised.* This *normalisation* process is essential to kernels that are sensitive to noises such as polynomial kernel. Assume $x \in \mathbb{R}^N$ is an input vector, the corresponding normalized vector $\widetilde{x}$ will be expressed as [23]:

$$\widetilde{x} = \frac{x}{\sqrt{\sum_i^N x_i^2}} \in \mathbb{R}^N \tag{2.14}$$

Features Space

To summarize, every data segment is characterized by a feature vector given by the tuple:

$$< P_{af3,alpha}, P_{af3,beta}, P_{af3,theta}, \; P_{af4,alpha}, P_{af4,beta}, P_{af4,theta},$$
$$P_{f7,alpha}, \; P_{f7,beta}, \; P_{f7,theta}, \; P_{f8,alpha}, \; P_{f8,beta}, \; P_{f8,theta} >$$

Where each $P_{c,b}$ is the *normalised* mean power of within EEG band b from channel c.

Cross-validation

The classification performance is measured by performing 10-Fold cross-validations for every scenario. The performance of the machine learning technique is categorised into:

- *Sensitivity*: test the effectiveness of the system in identifying *true positives*.

  $$sensitivity = true\ positives/(true\ positives + false\ negatives)$$

- *Specificity*: test the effectiveness of the system in correctly identifying *true negatives*.

  $$specificity = true\ negatives/(true\ negatives + false\ positives)$$

- *Error Rate*: test the ability of the system to correctly identifying any scenario - regardless of the target result:

$$error = (false\ negatives + false\ positives)/All\ cases$$

2.6.2 SVM Results

In the previous *statistical analysis*, the finding was the statistical figures of features representing differentiating process found from EEG signals are *strong*. Moreover, the *overall performance* of SVM in recognising those patterns have just further consolidate the original figures. The attained classification accuracies are relatively *high* and *uniform* in regards to individual subjects and the differentiating types. The classification results are also very consistent across kernel functions in use. The original intention of trying out SVM after the initial analysis was purely an attempt to compensate for the possibility of not attaining statistical significant figures from the *initial study*. However, not only do the SVM results back up the previous figures fittingly, they also indicate SVM's capabilities within an *online* BCI scenario. It has shown the maturity of machine learning algorithms in the research areas of BCI.

Radial Basis Kernel

Table 2.3 summaries the result of performing SVM with *Radial Basis Kernel*. The kernel was configured with value of σ set to *4*. By try-and-error, *4* is considered the optimal value for σ. For each test subject that were being studied, the classification accuracy is between *low 70* to *high 80* percent. The only significant under-performing instance is with test Subject 2 when he performed *task 2* - the error rate (in bold) was approximately 54% accurate together with a very low Specificity rate. The explanation for that can only be two reasons: the errors in EEG recordings for the particular scenario or, more likely a flaw in the experimental design. Similar results can also be obtained with other SVM kernel. Please refer to section 2.7 for more details on this experimental design flaw. One thing to notice is using RBF kernel results is the speed to train between the three kernels and is also significantly faster to train than with ANN. The average training time of each fold is about 0.05 second, quite fast for about 500 training samples.The RBF kernel also seems to fail to converge in instances when it is optimised with *Quadratic Programming* optimisation. Switching to *Sequential Minimal Optimisation(SMO)* method had addressed the issue.

	Subject 1	Subject 2	Subject 3	
Task 1	0.214	0.293	0.145	Error Rate
	0.757	0.722	0.896	Sensitivity
	0.815	0.692	0.815	Specificity
Task 2 (Shapes)	0.284	**0.459**	0.1193	Error Rate
	0.716	0.936	0.807	Sensitivity
	0.716	0.147	0.954	Specificity
Task 2 (Pictures)	0.249	0.282	0.164	Error Rate
	0.780	0.845	0.876	Sensitivity
	0.722	0.591	0.795	Specificity

Table 2.3: SVM classification results - Radial Basis Kernel

Polynomial Kernel

Table 2.4 shows the results of classification of SVM with a *polynomial kernel*. The polynomial order d is set to 2. From the table, one can tell this kernel yields the most *consistent* results overall. On rare occasions, the training failed to converge, especially with non-normalised data and it would take longer to train (about 0.19 seconds on average for each fold). SVM with polynomial kernel, however, achieves good *accuracy* when cross-validated.

	Subject 1	Subject 2	Subject 3	
Task 1	0.251	0.428	0.214	Error Rate
	0.819	0.722	0.915	Sensitivity
	0.680	0.422	0.656	Specificity
Task 2 (Shapes)	0.298	0.362	0.138	Error Rate
	0.780	0.734	0.789	Sensitivity
	0.624	0.541	0.936	Specificity
Task 2 (Pictures)	0.263	**0.388**	0.174	Error Rate
	0.849	0.807	0.869	Sensitivity
	0.625	0.417	0.784	Specificity

Table 2.4: SVM classification results - Polynomial Kernel

Multilayer Perceptron Kernel

Table 2.5 shows the results of classification of SVM with a *MLP kernel*. The results are *in line with* the results obtained from both polynomial and RBF kernels, with the exception from the Task 1, Subject 2 result (bold text in the table).

	Subject 1	Subject 2	Subject 3	
Task 1	0.276	**0.582**	0.255	Error Rate
	0.699	0.414	0.741	Sensitivity
	0.749	0.422	0.749	Specificity
Task 2 (Shapes)	0.349	0.294	0.243	Error Rate
	0.606	0.771	0.761	Sensitivity
	0.697	0.642	0.752	Specificity
Task 2 (Pictures)	0.324	0.396	0.298	Error Rate
	0.672	0.591	0.745	Sensitivity
	0.680	0.618	0.660	Specificity

Table 2.5: SVM classification results - MLP Kernel

Overall

In regards to the *overall* classification results, there some concern on the consistency in occurrence of *worst result* for each case. It appears that each kernel *under-performed* at various particular scenarios. However, it is quite inadequate to have an explanation for this because there are two factors that need to be addressed beforehand:

- Lack of test participants:
- Need to address certain experimental design flaws:

Besides that concern, the SVM classification results are quite consistent across the three kernels in general. Table 2.6 summarises the performance figure of all three kernels. The *Time* column is the *average* training time of each fold in *seconds*. The *bold* texts indicate the *best* values. SVM equipped with an RBF kernel achieved the best result out of the three. Not only does it perform *consistently better* than other options in term of accuracy, its training time is also very *desirable*. This kernel is the ideal kernel for SVM to be used in the proposed *online* BCI system.

Kernels	$\text{Error}_{(*)}$	$\text{Sensitivity}_{(*)}$	$\text{Specificity}_{(*)}$	Time
RBF	$\mathbf{0.246 \pm 0.103}$	$\mathbf{0.815 \pm 0.078}$	$\mathbf{0.694 \pm 0.229}$	0.0534 (s)
Polynomial	0.280 ± 0.099	0.809 ± 0.062	0.632 ± 0.165	0.1940 (s)
MLP	0.335 ± 0.104	0.667 ± 0.115	0.663 ± 0.103	**0.0490 (s)**

Table 2.6: Summarisation : SVM classification results
(*) $mean \pm standard\ deviation$

2.6.3 ANN Results

In other to have comparisons for the results obtained by SVM, an ANN classifier was used on the same dataset. The reasons behind that are:

- In the scenario the results are similar to the results obtained from SVM, it's possiple to confirm the *claim* that one could, computationally, identify the *differentiating process* patterns in EEG signals.
- An alternative in which SVM to be used as the classifier for the potential online BCI system.

Similar to those obtained with SVM, these classification results will be validated with 10-Fold cross-validation. Table 2.7 shows the cross-validation results of the ANN on the dataset. The training time was significantly longer than with the SVM, but the results seem to be very consistent across the test subjects and tasks. The results are also very close to the ones obtained from the SVM with the *polynomial kernel.* One would

	Subject 1	Subject 2	Subject 3	
Task 1	0.251	0.354	0.290	Error Rate
	0.718	0.536	0.625	Sensitivity
	0.780	0.755	0.795	Specificity
Task 2 (Shapes)	0.445	0.353	0.156	Error Rate
	0.385	0.651	0.853	Sensitivity
	0.725	0.642	0.835	Specificity
Task 2 (Pictures)	0.297	0.315	0.193	Error Rate
	0.707	0.598	0.764	Sensitivity
	0.699	0.772	0.849	Specificity

Table 2.7: ANN classification results

like to propose that the *similarities* he has identified in results using both classifying techniques had *statistically* underlined the *existence* of those *patterns* originally were set as the goal in this chapter.

2.6.4 Validating the Effects of Key-press on EEG Signals

As discussed in section 2.3.2, there exists a possibility that the *patterns* that have been identified with machine learning techniques above may be either:

- Influenced by the *sensory motor rhythm* (SMR) for the key-press actions since these and theoretical differentiating points are so close to each other.
- What successfully has been classified in previous analysis are indeed *SRM* that. Perhaps, one may end up achieving maybe just detecting the brain's *SMR*, not its *differentiation* signals.

In this section, however, the *refutation* to the above statements was demonstrated. It is ideal to perform a small analysis to determine if one could detect *any* motor sensory rhythm within the *same* EEG epochs. That leads to the hypothesis.

Hypothesis 2.2. If one *could not* effectively detect any motor-related rhythm within in the exact epochs that have been identified as the *differentiating* patterns in, he could claim that those *differentiating* patterns are not sourced or related to SMR.

Channel Selection

This time, *expert knowledge* was used as the method for *channel selection*. Previous work performed by neurologists such as Neuper [43] has shown that the neural activity which is related to executed motor movements is almost exclusively contained within channels C3, C4, and Cz of the EEG recordings. In regards to the EPOC EEG device, as it does not have those three EEG channels. In 10-20 coordinating standards, the closest alternatives to them that the device has are the four channels *T7, T8, FC5* and *FC6*. Hence, for the purpose of detecting SMR in this experiment, these four channels T7, T8, FC5 and FC6 will be used as the source channels. This is a nice *convenience* as this set of four channels are quite *independent* from the other set of four frontal channels used as the source to identify *differentiation* previously.

Signal Processing

The processing of EEG data is fact very similar to the previously mentioned method. The details are as follows:

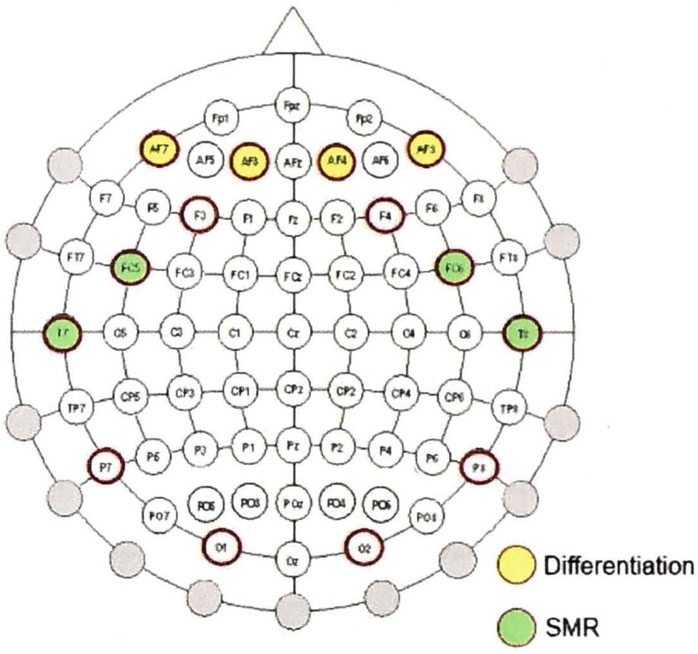


Figure 2.13: EEG channel selections: SMR vs. Differentiation

- EEG epochs are extracted as in the exact manner as done earlier:
  1. Range of [-0.5 -0.01] seconds at the event marker: considered as the epochs contain SMR: POSITIVE.
  2. Range of [0.01 0.5] seconds at the event marker: considered as the epochs do not contain SMR: NEGATIVE.
- For each epoch, FFt is used on the three channels: *FC5, FC6, T7* and *T8*.
- Each epoch is about 0.5 seconds long, which result in around 64 points at 128 sampling rate. Hence, the FFT window size is 64 (the next higher power of 2 to the value of data points of each epoch).
- Because FFT transformed data is symmetrical, the second half of the data was removed. This results in FFT *vectors* of 32 length in data points.

Feature Selection

In a very similar manner, every data segment is characterized by a feature vector given by the tuple:

$$< P_{fc5,alpha}, P_{fc5,beta}, P_{fc5,theta}, \quad P_{fc6,alpha}, P_{fc6,beta}, P_{fc6,theta},$$
$$P_{t7,alpha}, \quad P_{t7,beta}, \quad P_{t7,theta}, \quad P_{t8,alpha}, \quad P_{t8,beta}, \quad P_{t8,theta} \quad >$$

Where each $P_{c,b}$ is the *normalised* mean power of within EEG band b from channel c. Each input vector is labelled as either "1" for the POSITIVE cases or as "0" for the NEGATIVE cases.

Classification

For the purpose of classifying this dataset, Support Vector Machine (SVM) was chosen as the *classifier. Radio Basis Function* (RBF) kernel was also used as it helped produce the best overall results in previous classification tasks. To be consistent with previous tasks, the classification results are validated with 10-Fold cross-validation.

Result and Discussion

Table 2.8 shows the classification *error rate* of the dataset. From the table, the cross-validated correct rate is just around *0.50* on overall. For *two-class* classification, the results indicate not enough significant differences to separate the items within two groups. In other words, one could not effectively identify the SMR activities suspected to take place in parallel with the *differentiating activities* detected previously. The results support the claim that the *differentiating* activities identified previously was indeed:

- NOT influenced or in anyway the result of *SMR*.
- NOT the *SRM* as initially suggested.

Tasks	Subject 1	Subject 2	Subject 3
Task 1	0.500	0.500	0.485
Task 2 (Shapes)	0.490	0.505	0.482
Task 3 (Pictures)	0.504	0.500	0.496

Table 2.8: SVM classification Error rate - with channels T7, T8, FC5 and FC6

The above claim concludes this section. It suggests one could consider the *differentiating process* an *independent* and *separate* component to SMR. It further contributes to the claim of successfully *detecting* the *differentiating* process from studying EEG signals.

2.7 Lessons Learnt

2.7.1 Complexity of Differentiation Activities

Since the experiment conductors conducted the experiment and could observe how participants performed the differentiating tasks, it could come to the following realisation: *Most* of the differentiating tasks were too simple, so that each of them only took test subjects very little time to provision a differentiation. To make matters worse, each test subject became more and more familiar with each differentiating task as he progresses. Each participant seemed to be more adapted to the screen layouts as well as the contents on each screen after three to four screen into the series. As depicted by the shortened time in switching between screens, the effort in making the choices on later screens are notably *less* compared to the ones on the few initial screens. With that, there remain the following concerns:

EEG Hardware Sampling Rate

It may restrict the ability to capture the EEG signals of *differentiation* fully. Simple differentiating tasks were preferred and have been getting relatively short durations as the result. In the analysis, the signals within the 0.5 second period right before the marked event were considered that. With the sampling rate of 128 of the EEG device, there will be only around 64 samples to study the patterns of differentiation. Whether that number of samples is sufficient for the requirement is yet to be determined.

The Variety in Lenghts of Differentiation Period

As for the situation mentioned above, when a participant gets used to a certain type of task after a few tries, his *differentiation* will take shorter time to reach. How could one decide the ideal length for the time window to effectively cover as many differentiations as possible?

The Delays (or Tentative) in Differentiation

One participant may find the tasks at hand easier than the others have found, or he is just more acute to one type of task than to the other types. Nevertheless, those factors are very hard to quantify with such variety of differentiating activities that are also made by different people. Because of that, it considered appropriate to *bypass* those factors by analysing EEG signals with a *fixed length* time window. However, one have to concede that this approach is a bit *forced*.

Suggestions

One way to address the above concerns is to randomise the order of all differentiating tasks given to each participant - so that no consecutive tasks presented to a participant is of the same type. This would reduce the likelihood a participant to get accustom to a certain type of differentiation together with bringing up his alertness in making accurate and genuine differentiation. The main drawback for that approach is that it could result in many unforeseeable experimental variables.

More experimental scenarios should be designed so that the tasks require a higher cognitive level to solve. The longer that process takes, the better chance it will be recorded and identified correctly (harder choices, dynamically updated or animated, etc.).

2.7.2 Experiment Design

Lack of Room to Properly Identify Signals

This problem is best demonstrated by Figure 2.14. From the figure, one can see that in between *Screen 2* and *Screen 3*, the two windows reserved for Differentiation (positive) and NON-differentiation (Negative) cases are quite *close* to each other. If a test participant took *less* than a second to make one differentiation and then sent the key-press event (which is very possible), the half-second windows between those two screens will overlap each other. The result is different types of epochs could contain

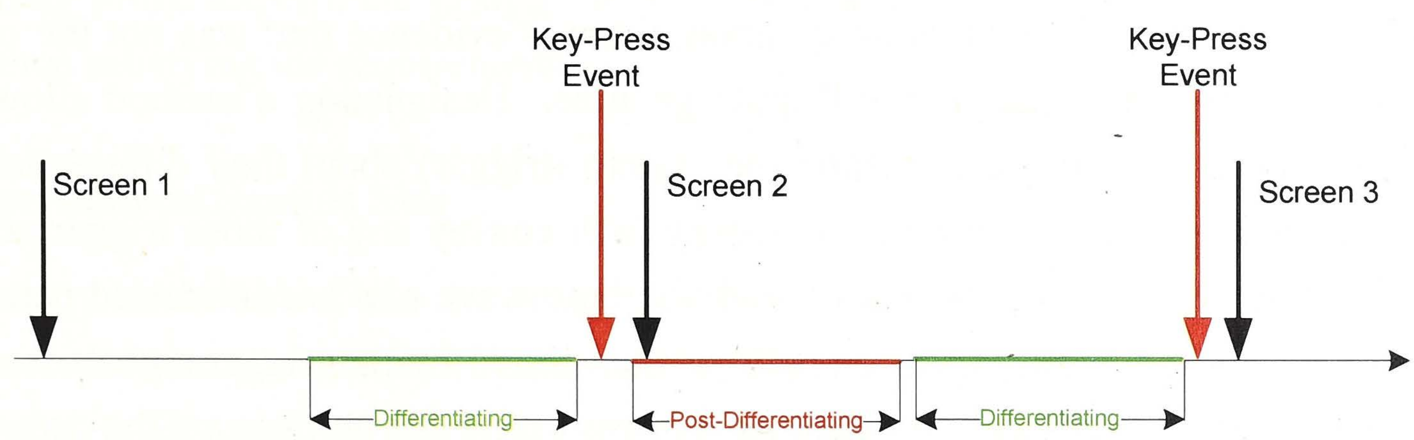


Figure 2.14: Different type of epochs between screens

the same parts of EEG signals within them. That would affect the effort of contrasting the differences in those epochs. Instances when test subjects took *less* than one second in between screens have been observed. Despite of that, those instances have been *temporarily ignored* in the analysis so far (i.e. every generated epochs were considered as *valid* and any suspected overlapping case was not eliminated). Because of that, a more effective mechanism to deal with these cases when generating these EEG epochs is *desirable*. In future trials, however, a simple modification to the experiment is *highly recommended*. It involves *adding delays* in displaying new screens after the key-press event of previous screens have occurred. Figure 2.15 demonstrate this solution. Figure 2.15 also suggests that the *length* of the added delays should be considered so that it is long enough to accommodate EEG signals that could be used as

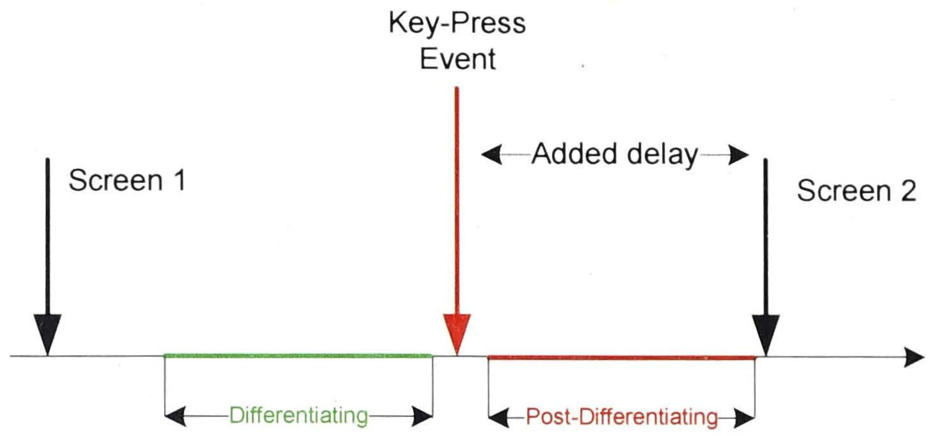


Figure 2.15: Added delay between screens

NON-differentiation signals. More than just a convenience, the suggestion *should* have been included in the original experiment as it is consideredthe proper way to allocate EEG signals for NON-differentiation epochs.

Utilise other Mechanisms to Notify Differentiation

The use of key-press as the mechanism to indicate *differentiation* could run the risk of the captured EEG signals of *differentiation* being susceptible to *sensory motor rhythm* (SMR). In section 2.6.4 of this dissertation, there is evidence that was not the case, this concern over this matter is still quite genuine. Designating a method allowing test participants to voluntarily inform the system (trigger) about their differentiation is a challenging task. Brainwaves, in theory, will convey any of those trigger activity (big or small), making the task of studying brainwave of some interested patterns without getting the fidelity of those patterns convoluted by the trigger signals is near impossible. Because of that, the ideal mechanisms fitting this purpose are the ones that involve:

- Human activities that have their EEG signals pattern well studied or established: so that their EEG signals could be easily be identified and isolated from the study.
- Human activities that require *little* or *insignificant* mental effort to perform. This is based on the assumption that activities that require *less* cognitive loads would introduce *less* noise to the overall EEG signals.

The key-press action was originally selected because it is considered to fall into both of the categories: SMR is a well-studied EEG rhythm and since the participant's finger is readily placed over the *key*, *little* movements of the finger are required to perform that action. Hypothetically, one could utilise another *equivalent* method for the same experiment. The benefits for that are:

- Firstly, one could compare its effectiveness to key-pressing mechanism in order to find the optimal trigger solution for the prospect of an *online* BCI system.

- Secondly, and more importantly, one could *contrast* the experimental results achieved with this new method against the ones obtained with key-pressing mechanism to genuinely *validate* the results achieved have so far, to see if the patterns identified in the previous part of this chapter was positively related to *differentiating*.

A More Robust Strategy for Marking Events

The majority of the EEG signals analysed in this chapter were extracted around the markers (offsets) of *one* type of event. Since those signals were only extracted in fixed-length windows (effectively 0.5 second before and 0.5 second after each marker). Perhaps, another strategy for placing event markers is recommended. This event marking strategy should be one that help identifying the start and end of the differentiating processes better than now. This section purely serves its purpose of explaining a shortcoming in the experiment design. It is yet to come up with such an event marking scheme satisfying the requirements.

Experimantal Sample Size

Contrary to the belief that the brain activities pattern associated with differentiation have been *identified* , it still requires to verify the results with more participants to support it. The analysis work was performed on EEG recordings of *five* participants, of which *three* results was used as the demonstration of the result. The other two participants' data was found *inconsistent* in term of quality of the captured data, probably due to improper subject preparation. Nevertheless, five participants is still on the small side for one to be statistically confident about the experiment findings. It is desirable to have at least ten participants for this experiment.

2.7.3 Data Analysis

The Possibility of a General Pattern in Differentiation

The main experiment design involves capturing the EEG data of participants performing multiple *mini* mental tasks throughout the session to have an insight on EEG patterns for those cases. So instead of just aiming for the goal of finding a *general* differentiation pattern from those scenarios, it would be sensible to take a step back and to study each of these mental tasks and the cognitive processes associate with them *individually*, at least for this early stage. So what have been done so far is just that: EEG data of each participant performing each individual task was studied, case-by-case. Only once attempts were made to *establish* a *general* patterns for multiple cases of *differentiation*, it failed: No significant enough evidence could be established for such *general* pattern. One of the more significant lessons learnt from studying those EEG data is the *singularity* aspect of it. It basically varies by participants and the differentiation tasks at hand. Perhaps the nature of EEG has simply denied that goal of identifying that general pattern. It still would be beneficial to revisit this problem with better approaches and methods.

The Variety in Durations of Differentiation

As far as EEG concerns, the captured signals are in time-domain. This means that if one to analyse the *differentiation* patterns, he should consider the signals in *between* different start and end points, depending on a number of factors. There is no guarantee that every *differentiation signals* is confined within that time window: One could take longer than 0.5 second, another could start before the start of the window while another differentiation could occur completely outside the window, etc. As mentioned before, The lack of a proper mechanism to "mark" the start and end of the differentiation have restricted to using fix-length windows. For this case, this is not necessary an ideal approach.

Study of Beta waves as an individual component

The peak values of delta, theta, alpha and beta wave signals were used as features for classification. However, only the beta waves are associated with attentiveness yet it is not known if that is best for decision making such as differentiation. So it would be ideal to have verifications if there is a connection between beta and differentiation in a more focused study.

Study of Task 3 Data

The study of Task 3 activities was not included in the discussion in section 2.2.3 because the two following difficulties have prevented from properly studying the EEG signals capture of the task:

- Unlike the activities in Task 1 and Task 2, the activities in Task 3 also involve the use of *previous* memories in making differentiation between scenarios. In short, they require participants, progressively, to form a conclusion from a series of scenarios. The problem faced here is that each participant will perform that task *differently* to the others (difference in techniques, methods, etc.). Those differences are quite difficult to quantify as they are very subjective to each person. They resulted in a number of *unknown* and *unidentified* factors will affect the outcomes of the analysis. However, those factors were yet to be addressed correctly during the course of analysing Task 3 data.

- The *initial* scenarios could dictate the final outcomes of the conclusions: Every scenario was introduced to participants in *random* order. This design was considered reasonable until it is realised that, since each participant start the task *differently*, they are not likely to conclude the task in a similar manner. That again, unfortunately, introduces more uncertainties into the analysis of the overall task.

- Furthermore, participants can conclude the task even before the scenarios end. The main issue with that is it is not clear about the time those conclusions have been made during the tasks. As differentiation made *before* the conclusions is expected to be different from the one made after the conclusions, this is a disadvantage because it was impossible to differentiate those two types from the data being collected.

A revisit to the Task 3 experimental designs to address the concerns above is recommended. From there, one can then go back and examine the data in a more confident manner.

Chapter 3

Implicit Differentiation and Reading

3.1 Introduction

This chapter demonstrates the study of Implicit differentiation activities during Reading tasks with brainwave captured with Electroencephalography (EEG). EEG, with its capability of recording brain-wave activities from the human scalp could exhibit potential in detecting series of differentiations made by humans while performing another activity. In this chapter, to prove the hypothesis, a BCI experiment was organised and proper method for effectively classifying EEG data in the scenario was introduced. Similar to Chapter 2, machine learning tools was also used to aid the study of brainwave data. The results consolidate the results of the previous chapter and further confirm the hypothesis that one could detect both Explicit and Implicit differentiation with EEG. In its conclusion, this chapter also discusses the potential of further use EEG for this work.

From this point onward till the end of chapter, any reference to *differentiation* will be considered as of the *implicit* type only.

3.2 Background

3.2.1 Reading and Implicit Differentiation

The foremost challenge any study of *implicit differentiation* has to face is designing the very experiment to detect it. If a subject does not become aware of performing such activities, how can he indicate their occurrence back to the researcher during the experiment. Even thought its existence is very *real*, it is still a very difficult concept to quantify, let alone to be detected via a measuring technique such as EEG. It is desirable for any experiment that deals with implicit differentiation to have the following attributes:

- **Uninformative** The test participants are not aware of the complete nature of the experiment in which they are participating. To avoid biases, it is ideal that each participant is not aware that he/she is being examined for differentiating/discriminating activities.

- **Diversion** Using other activities to "mask" any apparent discrimination to test subjects. It is to take their mind away from those activities in order to accomplish as natural discriminating activities as possible during the trials.

- **Verifiability** With those two above constraints, the experiment still has a valid way to verify these implicit activities in question.

In the experiment described in this chapter, reading task was used as the mechanism to hide any differentiation action from test participants. For most test participants, reading is an activity that is considered interesting or engaging enough to be able to diminish any awareness of these discriminations they make towards the contents they are reading. The reading contents would help differentiate the outcomes of the experiment.

3.2.2 Reading Activity

Reading is an activity that most human today perform on a very regular basis. They read and process information so much that reading skill becomes an almost second nature to them. The conjecture proposed here is based on that statement. So comprehensive and comfortable a human is in reading words, texts, that one would show, intentionally or not, certain behaviours that could be used to interpret his perception of the contents he reads. Understanding the meaning of words in sentences and paragraphs places a certain strain on a person's cognitive process. Depending on various contexts, such strain could go unnoticed by most of them. An example of such process would be, in order to comprehend a text, a person needs to build up linkages of information that he previously obtained with the current text. That cognitive process would be more significant if the text contains more information that he would, deliberately or not, associate back to. That assumption is reasonably correct because a normal person can only keep approximately seven pieces of information in their short-term memory [22].

There has been success in identifying such a process using gaze-tracking technique [51]. For this experiment, however, it is to validate the results with EEG. EEG is a brain-computer interaction technique that monitors brain-wave activities from the human scalp. It is suspected that the EEG signals would exhibit those aforementioned cognitive activities during reading tasks.

3.2.3 EEG and Eye-movements

A point for consideration in doing research with EEG is to deal with eye movement *artifacts* in EEG signals. Eye muscles produce considerable EEG signal noises and traditionally, EEG researchers would remove them from the signal analysis [25].

For this chapter, a different approach was proposed to that by *not eliminating* the effect of eye movements from the analysis. *Reading tasks* have one unique characteristic that supports this view: a person's eye movements tie quite strongly to their engagement to the contents being read. The increase/decrease in the amount of skipping forward and back-tracking activities found in the gaze *correlates* with the increase/decrease of the cognitive load in reading [51]. Studying of reading eliminating of eye-gaze noise could limit the potential outcome. As the aim is to identify the same link through the use of EEG instead of gaze-tracking technology, it would be convinient to take advantage of this *considered good* noise.

Another consideration is that the EEG signal, by nature, is *stochastic*. In regards to this experiment, it suggests that the 19 participants' EEG data should be processed and analysed *individually*. In this section, however, both ways were tried: considering each participant individually as well as *all* participants as a *whole*. The outcomes then were compared with each other. The results with the *gaze data* were also put against the one achieved from the previous experiment [51].

3.2.4 Methods

An experiment was conducted in which the test participants' EEG activities weew captured while they were performing reading tasks. A set of EEG features such as the frequency activations of EEG *alpha*, *theta* and *beta* bands etc. were chosen as the *factors* showing reader's *engagement level*. The link between level of engagement and the differentiating activities during reading reading is the key factor here: by showing different attention levels on different types of reading contents, one inadvertently shows the discrimination he makes on the items he's reading.

The experiment results are analysed for each individual participant against the whole set of participants. The aim is to verify the following hypothesis:

Hypothesis 3.1. For each participant, can one effectively identify the link between EEG signals and his level of engagement in the reading task? This confirms the presence of differentiation actions during the task.

In the event that one could confirm the hypothesis 4.1, it is likely that the following hypothesis can also be confirmed:

Hypothesis 3.2. Overall, can one achieve a general method to effectively identify the link between EEG signals and the level of engagement in the reading? This confirms the consistent presence of differentiation actions during those tasks.

For this chapter, together with confirming the 4.1 and 3.2 hypotheses , a method for processing EEG signals that is effective enough to be *considered* part of a *real-time BCI* system was also demonstrated. *Statistical machine learning techniques* were used to validate the findings.

3.3 Experiment

There are 19 participants for this experiment. They consist of 12 males and 7 males and all are within the 25-35 age bracket. None of them indicated of any known reading disorder. The experiment involves the participant reading some paragraphs from a computer screen while the computer captures their brain-wave activities via an EEG equipment. In total there were ten paragraphs for each participant to read. Similar to Chapter 4, seven of the paragraphs were taken from the paper "Keyboard before Head Tracking Depresses User Success in Remote Camera Control" by Zhu et al. [52]. The remaining three paragraphs were extracted from various sources about the same paper (miscellaneous paragraphs). Appendix A provides the images of the paragraphs.

Five of the paragraphs from the paper were chosen for the amount of useful information that was contained within and they are relevant to each other. The other five paragraphs (two from the aforementioned paper and the three miscellaneous ones) were chosen because of their generality and lack of specific technical information - they are irrelevant with the other five and also are irrelevant between themselves. Care was taken to make sure that this fact was not obvious to the experiment participants. These are the steps each participant has gone though:

- Finish going through the EEG set-up and calibration process

- Each participant reads ten paragraphs, one by one, on the screen.
- At any time, only one paragraph is displayed.
- Upon finishing reading one paragraph, say "Next" to indicate the experiment conductor to navigate to the next paragraph.

As mentioned before, the EEG reading of *each* participant on *each* paragraph was considered *independently*. The same condition were followed as close as possible: The cognitive activities of reading one paragraph *should not be influenced* by the cognitive activities found on reading *the other* paragraphs. Those effect should be limited as much as possible.

For that reason, the orders in which the ten paragraphs delivered to the participants were designed so that each participant would receive a *unique sequence* of paragraphs. The reason for the ordering is to avoid any common *cognitive trend* that may developed as two different users read the same *sequence of paragraphs*. Two participants, having read the first five paragraphs in the exact same order, may have lost interest in reading the remaining paragraphs in a similar manner/time. They may also develop similar *trends of thought* as they carry on. There is a possibility that their EEG readings contain unwanted patterns (or trends) that one might mistake for indications of reading cognition.

For each of the 19 volunteer participants, the general instructions are to read as if they were just reading any regular piece of text. They were also informed that they would not be questioned about the paragraphs they have just read read at the end of the trial. Figure 4.4(a) shows one of the paragraphs that each participant read.

```
In this research, we focus on importing computer
vision technology to undertake head tracking in
interface design for teleoperation activities.
The common remote control situation described
above is modelled by using a physical game
analogue: playing a table soccer game with two
handles. This has the advantage of being more
compelling for our student experimental subjects
than a more abstract task. We use student
experimental subjects as we have limited access
to the operators. We then propose a novel design
applying natural human head gestures for
controlling a Pan-Tilt-Zoom camera as an
effective approach to solve the camera control
problem.
```

Figure 3.1: An example of reading paragraphs

Similar to Chapter 4, the screen size of the monitor that the participants read on is a standard 19 inches. The resolution of the screen is 1280 by 1024. The screen is placed about 72 cm away from the participant face. The colour of the text is black on a white background. All paragraphs have the same font and occupy the whole screen.

For recording EEG signals, which are *considered* sensitive, special attention were paid to eliminate as much external distraction as possible during each trial:

- The head position of participants were secured with a *chin rest*. This is to minimise head/face movements (intentional or not) - which could greatly affect the EEG signals.

- The lighting in the room was dimmed to create an as relaxed environment as possible.

- The noise level of the room was limited to a minimum level.

- Possible sources of distractions (for instance mobile phones) were confined as much as possible.

The recording was *continuous* throughout the trial of each participant. After reading through 10 paragraphs (considered a trial) - the recording stopped and the captured data was stored in a file for analysis. Information of *timestamps* that mark the *start* and *end* of each participant's reading of each paragraph were also kept.

3.3.1 Hardware

There were two standard computers used in this experiment set-up, mainly for load-sharing purpose. They were standard PCs running the Windows operating system:

Visual computer

This computer has the responsibility for displaying the paragraphs on the screen.

- During each trial, experiment operator directly operated this computer to get started as well as changing paragraphs on the screen:

- This computer indicates those events to the *Capturing computer* by sending TCP/IP packets to the other computer.

- It also stores *miscellaneous* information related the visual responsibility that it performs into a database (This information was not in the end used for *analysis*)

Capturing computer

As its name suggests, this computer is connected to the EEG capture device and has the responsible for storing the EEG data captured from the device.

- Test operator performed initial test as well as calibration for each participant on this computer.
- It runs software that records the EEG signals from capturing device and stores them into the time series European Data Format (EDF) files.
- Incorporates the TCP/IP events notified by the *Visual computer* into the files as the recording session progressed.

EEG Device

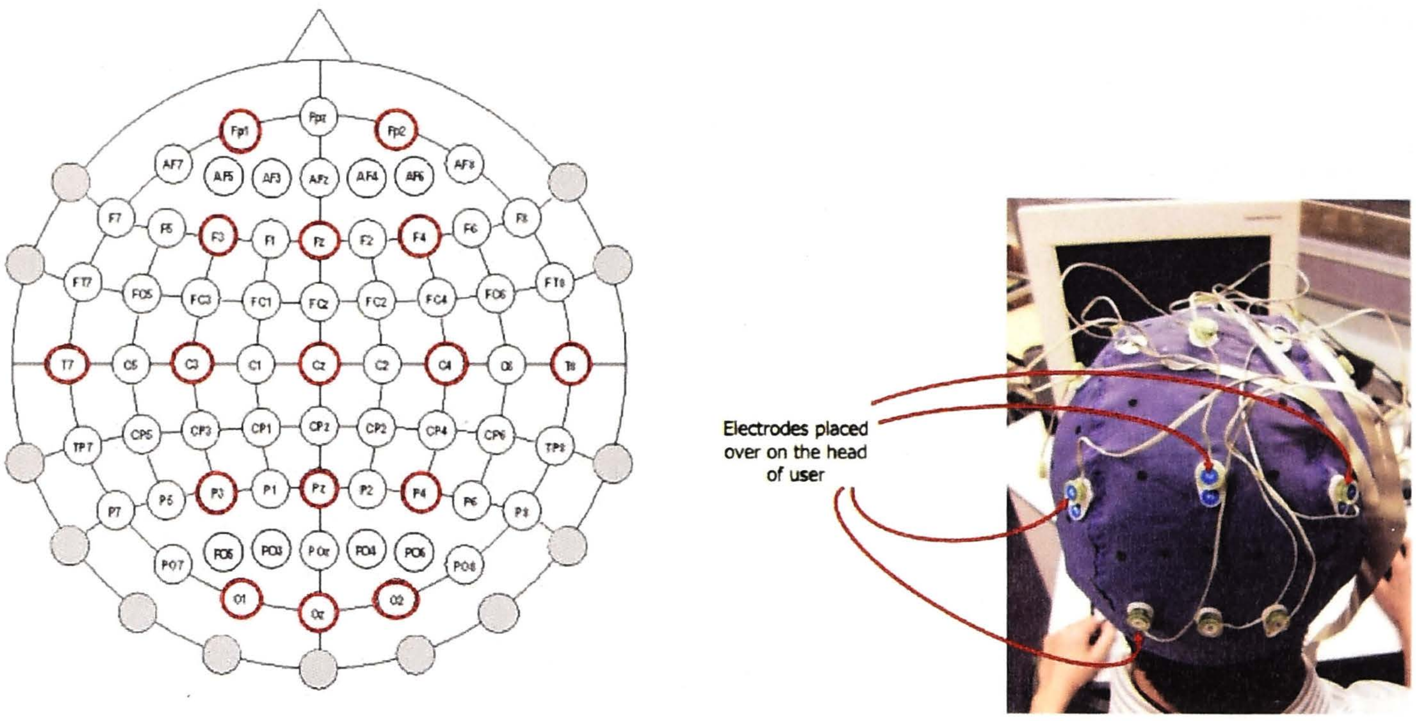
The EEG recording equipment used in this experiment is ActiveTwo provided by BioSemi:

- Supports up to 256+8 electrode + 7 event channels.
- Connects to the *Capturing Computer* via the Win32 driver.
- Hardware configured to capture at 16 kHz sample-rate.
- The electrodes set are located according to the 10-20 system.

For this experiment, BioSemi's first set of electrodes was used for recording, which included 16 electrodes and 2 ground channels . According to the 10-20 system, these 16 channels are: *Fp1, Fp2, F4, Fz, F3, T7, C3, Cz, C4, T8, P4, Pz, P3, O1, Oz, O2*. Figure 4.4(b) shows one participant set up with the 16 electrodes. Figure 3.2(a) shows the locations of the electrodes according to the 10-20 system, with the highlighted electrodes being used.

3.3.2 Software and Calibration

EEG devices require very careful *calibration*. The calibration process is required because it is essential that conductors know if *all* EEG electrodes have *good* contact to the participant's head scalp. By using the pin-type electrodes from BioSemi, it also required the use of electrode gel. The gel application job is almost a try-and-error process. There are a few factors that may affect the outcomes of this task: Thickness of hair, size and shape of the head and so on. The EEG set-up and calibration process can be summarised in the following steps:



(a) EEG electrodes locations of 10-20 system (b) Actual EEG electrodes layout

Figure 3.2: EEG electrode placements

- Place and secure the headcap over the participant's head.
- *Pump* electrode gel into the *electrode holders* provided on the headcap until the participant can "feel" that the gel has reached his scalp.
- Insert electrodes into the holders according to the labels (i.e. matching labels of electrodes and the electrode holders).
- Launch *Biosemi ActiView* software and start a live capture.
- Make sure all the electrodes can record *good* activity. i.e. the captured activities on all channels on display are correlated with each other in some simple tests such as eye-blinks and jaw movements, etc.
- To identify *bad* electrodes, study the Electrode Offset tab in ActiView and look for channels with abnormal offset compared to the rest of channels.
- Correct the electrodes (readjusting and providing more gel, etc.) until no *bad* electrodes can be found.

Figure 3.3 demonstrates the relationships between the software were used in this experiment. They are as follows:

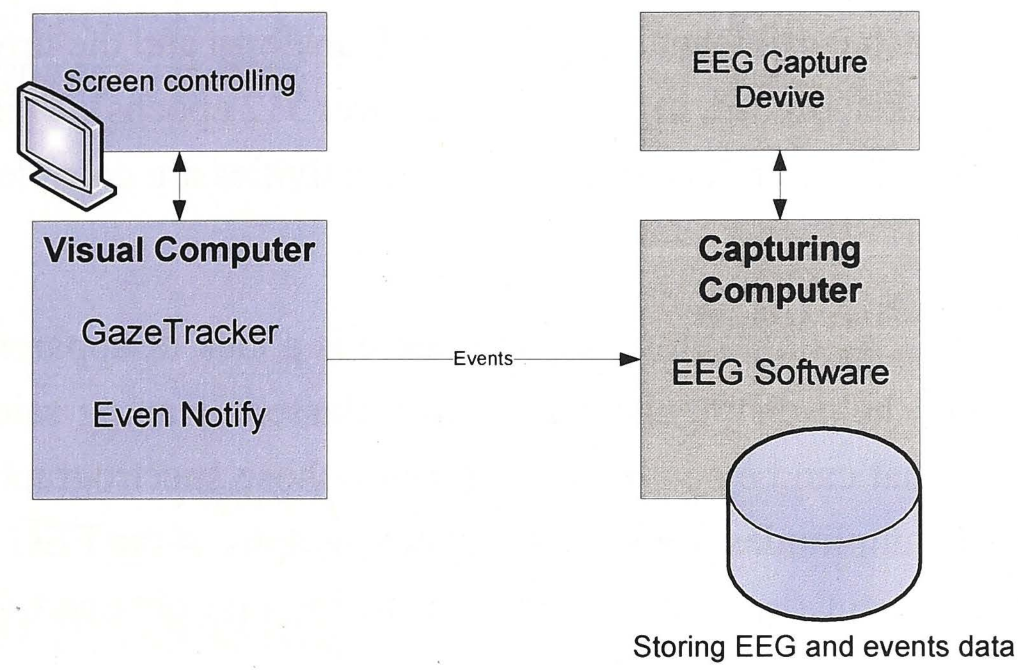


Figure 3.3: EEG computers and the software set-up

BioSemi ActiView

Runs on the *Capturing Computer*. It is the capturing software provided for use with BioSemi AciveTwo. However, since it does not meet some of the requirements, conductor only used this tool for the *calibration step*.

In-house EEG capturing application

Runs on the *Capturing Computer* for the purpose of recording the captured EEG data. This software was built on top of the API provided by BioSemi. The captured EEG signal is stored into EDF files. This software can receive and recognise custom *TCP/IP events* sent from the *Event Notifier*.

In-house Event Notifier application

A light-weight tool that runs on the *Visual Computer*. As it resides on this computer, it can detect interactions (keyboard inputs, mouse), process them and then notify the *EEG capture software* as *TCP/IP events*.

3.4 Preliminary Analysis

For the initial analysis, time-frequency analysis was ran on the raw EEG data, as suggested by Makeig [39]. One hope to observe the differences in time-frequency distributions of EEG signals captured from a person reading a relevant against irrelevant piece of text (English paragraphs). This analysis was performed on the first raw EEG channel (Fp1) of each participant using Fourier Transform and the time resolution was going to be (time taken to read, in milliseconds, over 512 epochs). This is be consistent with the findings in Chapter 2, where the frontal activities are considered to be related with differentiation

The initial observation has revealed that there is a lack of apparent and consistent features that could help distinguishing the two classes. Having said that, there is a minor difference that can be spotted by observing those spectrographs - that is, there are more drops in amplitudes found on the spectrographs of the EEG signals recorded from reading irrelevant paragraphs compared to the one obtained from reading the relevant paragraphs.

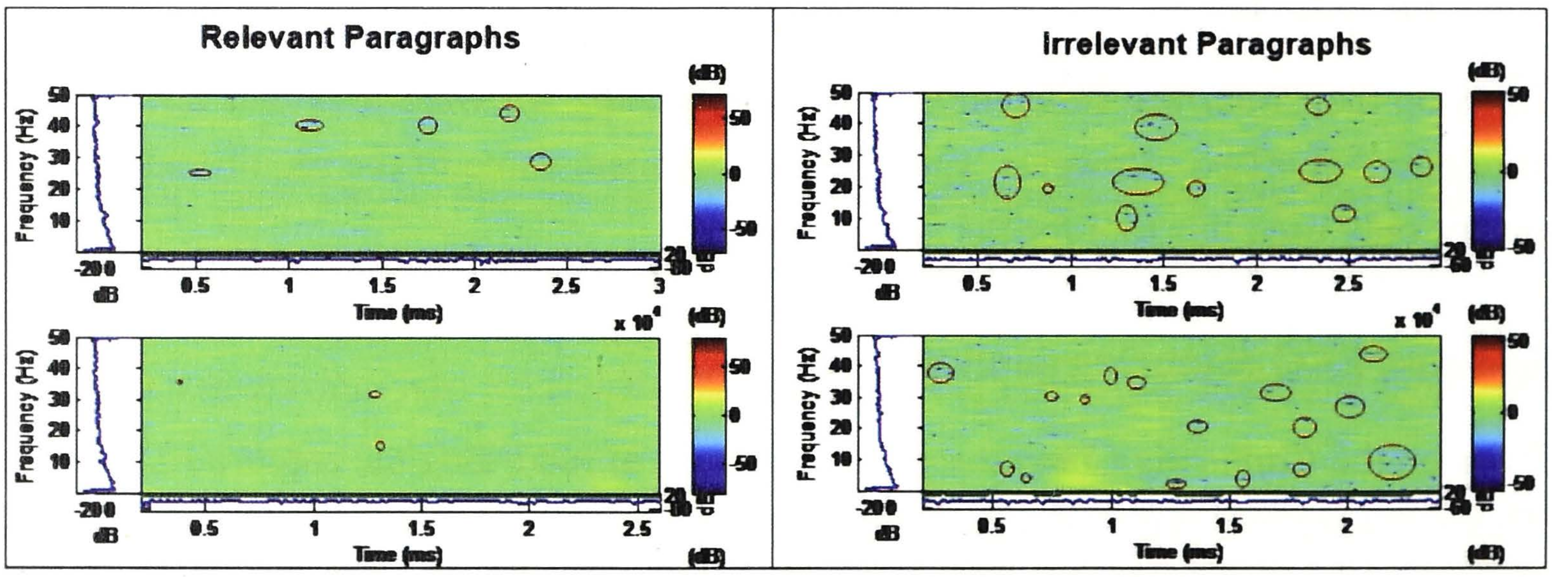


Figure 3.4: Visualisation of EEG readings activity of paragraphs

Figure 3.4 demonstrates the above observation. It includes spectrograms generated by analysing the EEG signal a participant reading two relevant paragraphs (Left-hand Side) and two irrelevant paragraphs(Right-hand side). The circled spots are some of the locations where the drops in amplitude can be identified.

The initial analysis has shown that there is a possibility that, using statistical machine learning techniques, one could effectively classify the two cases. For that purpose, a standard Artificial Neural Network configuration was used as the foundation. Optimization will be considered if the initial results are promising.

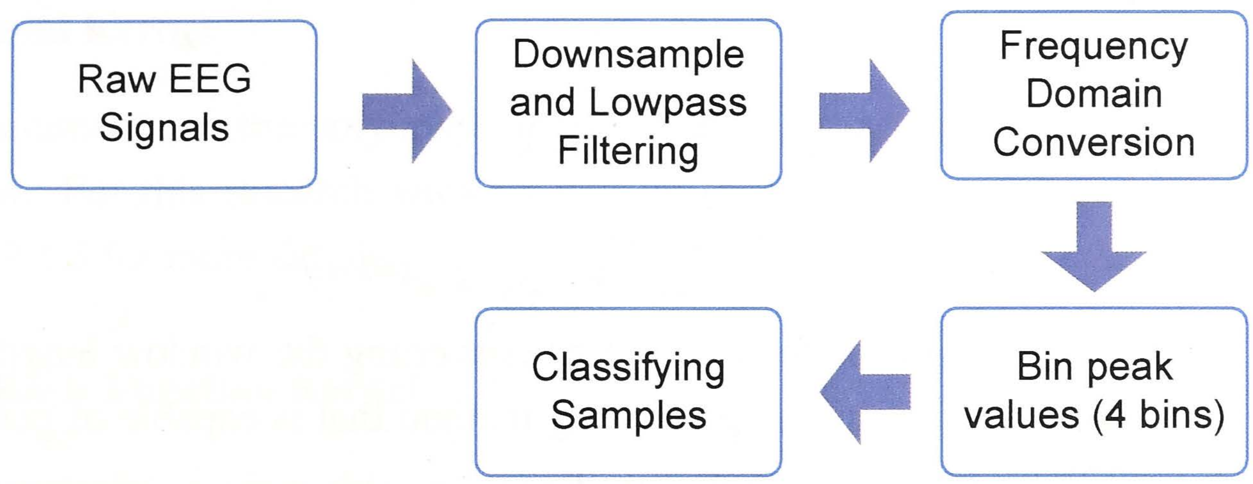


Figure 3.5: Signal processing process

3.5 Signal Processing

3.5.1 Fast Fourier Transformation

The EEG signal was originally collected at 1024 Hz, which were down-sampled to a rate of 256 samples per second. A low-pass filter of 60Hz was ran to eliminate the unwanted EEG frequencies. The time-domain EEG signal was then broken into epochs (windows) of one second. Fourier-Transformation (FFT) was performed on each window, for each of the 16 channels. The data from the FFT was binned into four frequency ranges: Delta, (0-4Hz), Theta (4-8Hz), Alpha (8-13Hz) and Beta(13-30Hz). Figure 3.6 shows the output of the FFT with the colour-coded ranges.

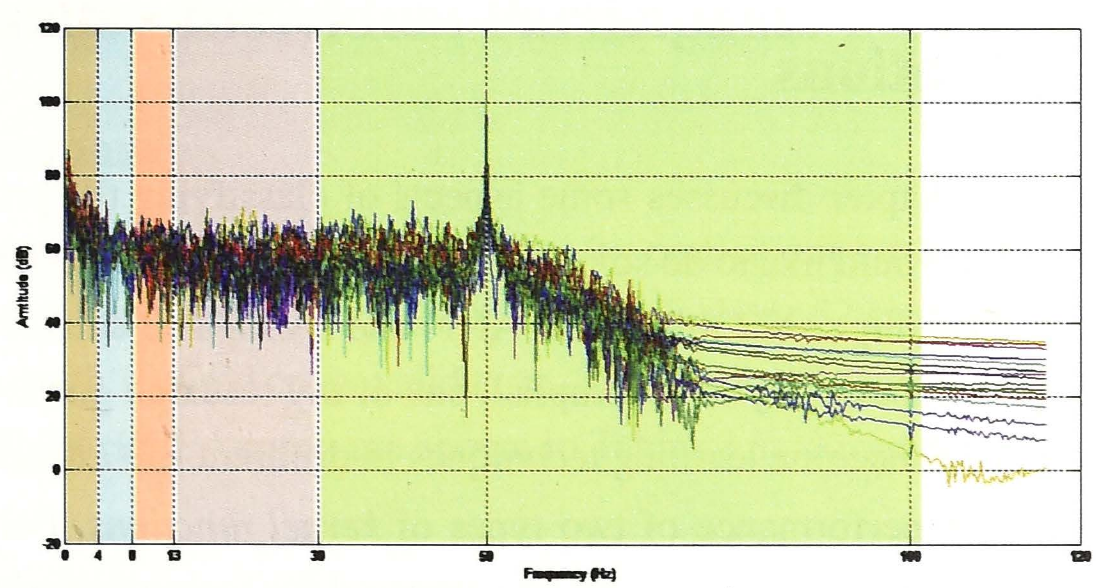


Figure 3.6: Visualisation of an FFT transformed for 16 channels of an epoch

FFT was chosen because it is a very efficient transformation algorithm. The implementation of the FFT is given based on the Cooley-Tukey [13] approach. Optimization of window size could further improve the performance of this model. The Discrete Fourier transformation of a vector x of N length is given by:

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)} \tag{3.1}$$

The Nth root of unity ω_N can be calculated by $\omega_N = e^{(-2\pi i)/N}$.

FFT is efficient enough for the purpose of converting the window length of one second. Using it results in a signal processing method that is capable of performing live (real-time) EEG signal processing.

3.5.2 Peak Detector and Feature Set

A simple peak detector (sensitivity Δ of 0.7) runs on the FFT bins to detect peak values for each EEG band. This detector considers a *peak* to be the highest point between the *valleys* of the value series. The implementation is a straight-forward linear search and compare to identify local peak values on vectors.

Since the activities of the four EEG bands (Alpha, Beta, Theta and Gamma) were interested in, a window (epoch) will have four features representing it. Since there are 16 channels, a regular dataset will have 64 features (4 x 16). The same process 3.5 will be applied on the subsequence epochs until the end of the EEG recording. The samples are labelled "1" to indicate reading relevant text against "0" for reading irrelevant text.

3.6 Classifications

This section of the chapter discusses some aspects of classifying the processed EEG data. Furthermore, I would like to do some comparisons between two type of statistical machine learning techniques for this work: ANN and SVM. The dataset produced so far provides a good opportunity to accomplish one of my research goals: To critically evacuate the usage of machine learning techniques in different EEG datasets.

As for SVM, the performance of two types of *kernel functions*, Polynomial and Radial Basis Function (RBF) was also being compared for their popularity with usage in classifying EEG data.

3.6.1 Support Vector Machine

Please refer to Section 2.1.5 for more information related on this classifier. The specific parameter settings for this work are as follows:

Polynomial Kernel

The d parameter, or the polynomial order, allows one to customize the feature conjunctions. For this research work, d will be set to 3 (cubic polynomial). Refer to Section 2.1.5 for more details.

Radial Basis Function Kernel

The σ parameter is adjustable and would dictate the performance of the kernel. With this dataset, the optimal σ value has been found to be 5. Refer to Section 2.1.5 for more details.

3.6.2 Artificial Neural Network

With the dataset, it is suggested that a standard ANN configurationwould be sufficient. The ANN setup constructed for this experiment is a feed-forward, back-propagation network.

This network has one hidden layer containing 20 hidden neurons and one output. As for the neural network optimization algorithm, the Levenberg-Marquardt optimization (ML) training algorithm was used. Please refer to Section 2.1.6 for more information related on this classifier.

3.7 Results - Individual Participants

3.7.1 Summary

The dataset was devided into smaller groups by participant - called P1, P2 all the way to P19. The average sample size of reach group is about 57 samples. For each group, ANN was ran with 10-Fold cross-validation. The results of the ANN, SVM(Polynomial kernel) and SVM (RBF kernel) are shown in Tables 3.1, 3.2 and 3.3 respectively.

3.7.2 Discussion

Accuracy

The results displayed on the tables above demonstrate the effectiveness of the chosen method (data processing and classifications). In terms of accuracy, one can see from Figure 3.7 that the classification accuracy is consistent for the three methods. The average accuracy rate is about 95 percent, which is quite encouraging for the task of

	Accuracy	Specificity	Sensitivity		Accuracy	Specificity	Sensitivity
P1	0.928	0.966	0.872	**P11**	1.000	1.000	1.000
P2	0.918	0.891	0.963	**P12**	0.927	1.000	0.833
P3	0.986	0.977	1.000	**P13**	0.968	1.000	0.920
P4	0.924	0.921	0.929	**P14**	1.000	1.000	1.000
P5	0.781	0.667	0.929	**P15**	1.000	1.000	1.000
P6	0.931	0.889	1.000	**P16**	0.972	1.000	0.941
P7	0.984	1.000	0.958	**P17**	0.966	0.962	0.972
P8	1.000	1.000	1.000	**P18**	1.000	1.000	1.000
P9	0.900	0.871	0.947	**P19**	0.985	1.000	0.963
P10	0.979	1.000	0.944	**Avg**	0.955	0.955	0.956

Table 3.1: ANN classification results for 19 participants (individual)

	Accuracy	Specificity	Sensitivity		Accuracy	Specificity	Sensitivity
P1	0.876	0.948	0.769	**P11**	0.906	0.948	0.737
P2	0.959	1.000	0.889	**P12**	0.927	1.000	0.833
P3	0.973	1.000	0.933	**P13**	0.984	1.000	0.960
P4	0.970	0.947	1.000	**P14**	0.980	0.947	0.955
P5	0.719	0.667	0.786	**P15**	1.000	0.667	1.000
P6	0.966	1.000	0.909	**P16**	0.861	1.000	0.706
P7	0.951	1.000	0.875	**P17**	0.966	1.000	0.917
P8	0.981	1.000	0.957	**P18**	0.967	1.000	0.926
P9	0.960	1.000	0.895	**P19**	0.941	1.000	0.852
P10	0.938	1.000	0.833	**Avg**	0.938	0.977	0.881

Table 3.2: SVM classification results for 19 participants (individual) - Polynomial Kernel

classifying EEG signals. There is still minor inconsistency in the achieved results - with P5 achieving about high 70% accuracies (Table 3.1). It indicates that further studies could be done to investigate the profiles of these participants.

The SVM accuracy depends a lot on the adjustable parameters. In the case of the RBF kernel, if σ is any number lest than 4, the results are not as good. In fact, those SVM kernel parameters were chosen from some try-and-error process. For that, one can confirm that they are considered ideal settings for this dataset.

In relation to a real-time system for predicting this kind of scenario, it is shown that even with a relatively small effort of training i.e. reading tasks of only 10 paragraphs,

	Accuracy	Specificity	Sensitivity		Accuracy	Specificity	Sensitivity
P1	0.938	1.000	0.846	**P11**	0.925	1.000	0.789
P2	0.973	1.000	0.926	**P12**	0.951	1.000	0.889
P3	0.959	1.000	0.900	**P13**	1.000	1.000	1.000
P4	0.985	1.000	0.964	**P14**	1.000	1.000	1.000
P5	0.875	0.944	0.786	**P15**	0.956	0.944	0.895
P6	0.966	1.000	0.909	**P16**	1.000	1.000	1.000
P7	0.984	1.000	0.958	**P17**	1.000	1.000	1.000
P8	0.981	1.000	0.957	**P18**	0.967	1.000	0.926
P9	0.960	1.000	0.895	**P19**	0.985	1.000	0.963
P10	0.958	1.000	0.889	**Avg**	0.966	0.997	0.921

Table 3.3: SVM classification results for 19 participants (individual) - RBF Kernel

one can still achieve a quite successful classification result. The potential for a working system based around this experiment is very promising.

Performance

With the classification accuracy among the three methods are quite similar, the performance of the three, on the other hand, is quite different. The training time of each scenario (begin and right after the training finishes) was measured and has been found that, ANN was the worst performer so far. It took about *920 seconds* to finish (19 participants x 10 folds each) - that's on average about 4.8 seconds per fold. The difference between that performance figure compared to the two SVM-based approaches is quite significant. The SVM with the *polynomial kernel* took just about *9.9 seconds* (0.05 seconds / fold) while the SVM with the *RBF kernel* was the quickest to converge with **3.2 seconds** (average 0.01 seconds).

In terms of speed, the SVM-based approaches held a big advantage over the ANN based solution. This EEG dataset seems to favour the SVM more than it does with ANN. In this particular scenario, the MBF kernel also seems to be the more suitable option than the Polynomial kernel thanks to its very fast training speed. Generally speaking, SVM seems to be the more fitting tool for this particular EEG problem - or EEG in general.

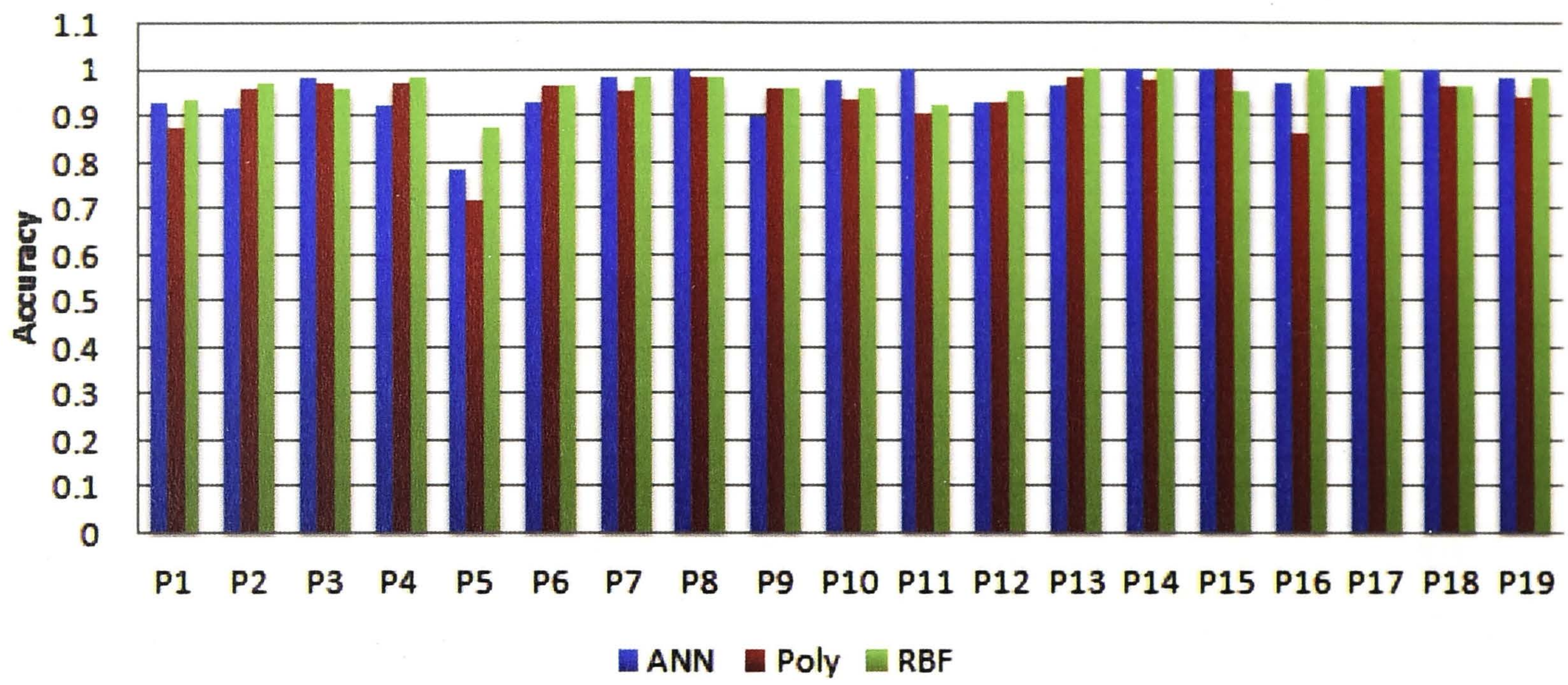


Figure 3.7: Comparison of classification accuracy of methods

3.8 Results - All Participants

This section explains the preliminary attempts to identify an EEG patterns of differentiation for this scenario regardless to the individuals from whom the EEG signal was collected. To serve this purpose, the combined dataset from all participants were classified with the same ANN set-up. This is to confirm the hypothesis that there is a *general* EEG pattern. Similarly, the results were validated with 10-Fold cross validation. The results are shown in Table 4.3.

Table 3.4: ANN classification results for 19 participants (together)

Method	Accuracy	Specificity	Sensitivity
EEG - ANN	**0.817**	0.889	0.716

The results how that the overall dataset provides a lower accuracy than the individual dataset. This is expected because EEG signals are subject to differences in a number of aspects (time, person to person, mood, etc.). But one can still observe that the overall classification accuracy obtained is almost on par with the one obtained individually. This is indeed very encouraging as it suggested that the existence of a general pattern of differentiation activities during reading is likely.

The result obtained here also suggests that by increasing the amount of training together with further optimisation of ANN configurations, one could improve it to a more desirable level . It has laid the foundation for further research work in this area.

3.9 Summary

This summaries the investigation of implicit differentiation with EEG - in perspectives of Signal processing and Classification using machine learning techniques. The results demonstrate the capability of using EEG and statistical machine learning classifiers to detect these activities in a very effective manner. The experiment results give a strong case in confirming the originally stated hypotheses. The results are also specific to the reading task which suggests that further work should be done to really achieve the goal of detecting these activities in more general scenarios. Future experiments on this subject should not limit themselves to reading while other more general activities are recommended to be tried.

This work has been published into the proceeding of *International Conference on Neural Information Processing 2011* conference held in Shanghai, 2011.

Chapter 4

Implicit Differentiation with Gaze-tracking

4.1 Introduction

The behaviours of a person's eye-gaze can reveal a wealth of information about him. For that, gaze-tracking is considered to be a capable alternative to EEG in detection of implicit differentiating activities. This chapter describes the investigation on relationships between gaze-tracking and EEG.

Experiments in reading tasks equipped with gaze-tracking were organised for over thirty participants and the experimental outputs were classified with Artificial Neural Networks (ANN) as well as Support Vector Machine (SVM). Both tools provided results with an approximately 80 percent accuracy. This outcome support the initial assumptions, proving that gaze-tracking is just as effective as EEG in studying these cognitive activities.

"The eyes are the window of the soul." (English Proverb)

4.2 Background

4.2.1 Eye-gaze and Reading

Everyone is taught to read (in English) the same way: Read a line from left to right and then drop down to the next line once the end of the current line is reached. As a beginner, one followed this simple rule very closely but as he gets more adept at reading English text, it is no longer the case. People develop their own personal behaviours when reading, and that they probably do not notice they even do so. One aim of this research is to characterise these behaviours and to introduce a meaningful model that can show how a person is engaged in the reading activity.

In regards to reading of English text, how much can one learn from a person's gaze pattern? It is known that while reading, he inadvertently form rational connections between pieces of information he picks up from the text. That reflects in certain disruptions in the norms of reading and that gives clues to the interest level in reading activities.

One could generally categorise the eye behaviours during reading tasks into three groups:

- **Knowledgeable Movement** describes small movements after fixations. In this situation, one could roughly say that this person is reading *word-by-word*. This would be considered *normal reading behaviour* as it is considered that this pattern occurs most often. In this state of mind, a reader can *foretell* what his next target is, which is somewhere around the last fixation.

- **Searching** It is considered that this pattern to be a more *extreme* version of the above patterns. The shift in viewing angle is considered to be around 3° from the last fixation point. An example for this pattern is when a reader kips to the end or beginning of a sentence, or to a different line within a paragraph. This is when a person knows the general area of their target is and they quickly move to there and locate it.

- **Unordered searching** is when a person has made too many *saccades* (movements between fixations) since the last fixation in order to know where they should be looking. This behaviour could mean that the user is searching for their next target or they are simply thinking.

So what would one get from studying those above patterns? It is quite safe to assume that the cognitive loads required vary from pattern to pattern - and if one would gather enough statistical information of each one (frequencies, instances, etc.....), he could potentially establish a reading profile for him, then use it to measure the engagement level of a person on the piece of information he is reading.

4.2.2 Reading Comprehension

Comprehending the meaning of words in sentences and paragraphs is indeed a great (unnoticed) strain on a person's cognitive process. In order to comprehend text a person needs to be able to read quickly because a person can, generally, only keep seven pieces of information (± 2) in their short-term memory [22]. Any additional information is quickly lost and cannot be recalled. This general rule stands for many different kinds of information from the very simple (letters, or words) to the very complex (sentences, paragraphs). This allows a fluent reader to be able to *chunk* related information together so that they can get more words into their short term memories. The above phenomenon results in certain disruptions in reading patterns. It is believed that these stochastic behaviours are the keys to effectively quantify the reading engagement level of a person. Previously in 2009, a study was organised in the research group to investigate using eye-tracking to analyse reading behaviour. Even still in a preliminary stage, it showed the potential in using machine learning approaches to classify eye gaze patterns [15]. So, the purpose of the research behind this chapter is to consolidate the previous studies' results and to propose a feasible model for classifying gaze-pattern with machine learning methods such as Neural Networks. As for that, another experiment was conducted in which conductors captured test participants' gaze activities while they were performing reading tasks. Then a set of features of those data: reading time, fixations time, differences in X and Y coordinates, etc.... was analysed to identify the key factors to indicate user engagement level in reading. With the positive result obtained from that, a simple but effective approach to measure a person's interest in the materials he is reading was proposed.

The original goal was to investigate if one could associate certain gaze behaviours to various cognitive tasks. In this chapter, the sole focus is on the *reading task*. The hypothesis that are being considered is as follows:

Hypothesis 4.1. By capturing people's gaze while they are reading, one would be able to tell if test subjects are interested in or how carefully they read the contents on the screen. This confirms the presence of differentiation actions during reading.

This work's aim is to validate the above statement. It also proposes a novel method for detecting the level of engagement in reading based on a person's gaze-pattern. The method in question is based on the previously mentioned gaze features. It involves three design principles to make it a lightweight yet effective method for this purpose. one of them is the use of Artificial Neural Network (ANN) technique to validate the effectiveness of the proposed approach. By combing this solution with a reasonably simple ANN, one could introduce a very achievable real-time system. This solution is also flexible in combining with other classifying techniques. This further strengthen the above claim by achieving a comparably accurate result with Support Vector Machine (SVM).

4.3 Experiment

4.3.1 Background

35 participants was involved in this experiment. The experiment involved the participant reading some paragraphs from a computer screen while the computer gathered their eyes(gaze) movements with gaze-tracking equipment. There were two variation in the way participants perform the task, which will be further discussed below:

4.3.2 Participants

The 35 volunteered participants were divided them into two group, called group A and Group B. Group A contained 13 people while group B had 22. The difference between the two groups is the amount of steps it takes for each participants to complete the tests, and the instructions they are given.

Group A

Group A were people that had been informed that they would have to answer questions about the paragraphs they read toward the end of the experiment session. In more details, there were three steps that each participant had to perform:

1. **Reading the Paragraphs** The first part to the experiment was the reading of the paragraphs. It involved participant reading some paragraphs from a computer screen while a gaze capturing device was "watching" their eyes:

    - The amount of paragraphs was *ten* for each participant.
    - That one set of paragraphs was used for every participant.
    - The order of them appearing to each participant was designed to be different for every each of them
    - The participants needed to say "Next" to indicate the conductor to move to the next paragraph in the sequence.

2. **Reading and Answering the Questions** Right after reading the paragraphs, each participant in this group was asked to answer five *multiple choice questions* on the paragraphs.

    - Similarly to the paragraphs, the questions were presented to him on the computer screen together with the multiple choice answer. His eye gaze was recorded as he read the questions as well as the answers.

- To choose an answer he thought was the best, the participant just needed to say the letter associated with that answer. The conductor of the experiment then recorded the answer and moved to the next question.

3. **Describing the Paper and Ordering the Paragraphs** This step of the experiment is designed to evaluate a participant's understanding of the material he just read:

   - This part of the experiment was conducted on paper and no eye tracking was used.
   - Each participant was asked to write about the general topic in *one sentence*.
   - Each participant was given the printed version of the paragraphs he was reading on the first part. The sheets were organised in the same order they appeared on screen beforehand.
   - The participant ranked each paragraph on the scale of 10: the one with the most useful information for completing the questions, as number "1", and the one with the least information, as "10". As there were 10 paragraphs, each paragraph would have a distinctive score eventually. He could reread the paragraphs as much as he needed during this step.

Group B

Participants within group B only had one step in common with those in group A - that is the first *Reading the Paragraphs* task. Hence the experiment of people in group B can be described can be described as follows:

- This involved the participant reading some paragraphs from a computer screen while the computer was watching their eyes.
- The number of paragraphs are ten for each participant.
- That one set of paragraphs was used for every participant.
- The order of them appearing to each participant was designed to be different for each of them
- The participants would say "Next" to indicate the conductor to move to the next paragraph in the sequence.

Another difference with the participants in group B compared to those in group A is that they were *informed* that there would be no further task to perform after reading, and that they would not be asked questions about the paragraphs they just read. In fact, they were recommended to *read as naturally as possible*.

It is also worth mentioning that the experiment performed in chaper 3 is the same experiment that involved the participants of Group B. That is the participants in Group B will also have thier EEG readings recordered in parallel with thier gaze-tracking readings.

4.3.3 The Paragraphs

In total there were ten paragraphs for the participants to read. Seven of the paragraphs were taken from the paper "Keyboard before Head Tracking Depresses User Success in Remote Camera Control" by Zhu et al. [52]. The remaining three paragraphs were extracted from various sources (miscellaneous paragraphs). Five of the paragraphs from the paper were chosen for the amount of useful information that was contained within. The other two paragraphs from the paper and the three miscellaneous paragraphs describing that paper were chosen because of their generality and lack of specific technical information. Some care was taken to make sure that this fact was not obvious. Appendix A provides the images of the paragraphs. The paragraphs were presented to participants in different orders to prevent any specific paragraph ordering affecting the results. This was an experiment design choice to help showing which participants could look at the bigger picture even when the information was out of sequence and scattered.

4.3.4 Equipment Setups

During the experiments the participants read all the paragraphs off a screen which was connected to the same computer that was recording their eye movement. The computer was a standard desktop machine that was running on Windows XP with SP3. See Figure 4.1 for more details.

- The computer has two screens connected to it, one for controlling and monitoring the experiment and a 19 inch screen for the participants to read the paragraphs and questions off.
- The screen that the participant were reading off was an LCD and was set to a resolution of 1280 by 1024. All the paragraphs and questions were set to the same resolution so no scaling was required.

- To assist the gaze tracking system (by stabilizing the participant head positions), a chin rest was fixed in front of the screen. It was 72 cm away from the centre of the screen. The centre of the screen and the chin rest were both 100 cm off the ground. This height of it was fixed for all participants. Conductor adjusted the height of participant's chair according to their height instead. Please refer to Figure 4.2 for more details.

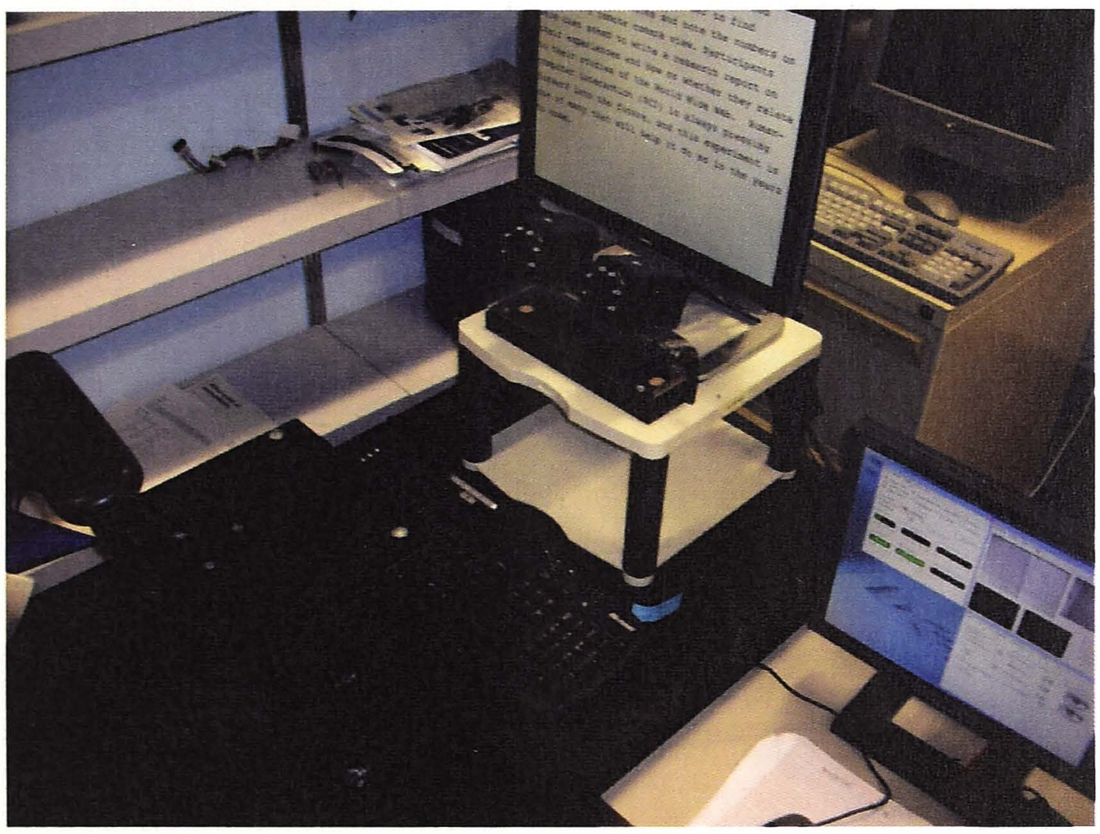


Figure 4.1: The setup of the experiment

The eye tracking system that was connected to the computer was FaceLab 4.5 provided by Seeing Machines. The system consisted of two Sony VFCB-EX480B cameras and infrared depth sensor. The tracking software in used was Seeing Machines Face-Lab 4.5:

- The two cameras were located at the bottom of the screen that the participant was reading from and they were mounted 9 cm in front of the screen with a distance of 63 cm to the chin rest. Figure 4.2.

- The cameras were mounted 14 cm apart, each one was 7 cm away from the vertical centre of the reading screen. At the centre of the gap between the two cameras was an infrared light source.

- Before the experiment could begin, the system was calibrated for each participant. This included manually setting the head model in Face-Lab and checking the accuracy of the gaze calibration.

- Six participants were removed because of the poor accuracy of the gaze calibration.

- The Seeing Machines Face-Lab software recognised the gaze points and forward those data via a TCP/IP port (i.e. as TCP packets). It is configured this way because it would be more convenient for another software (see Figure 4.3) to store the gaze data.

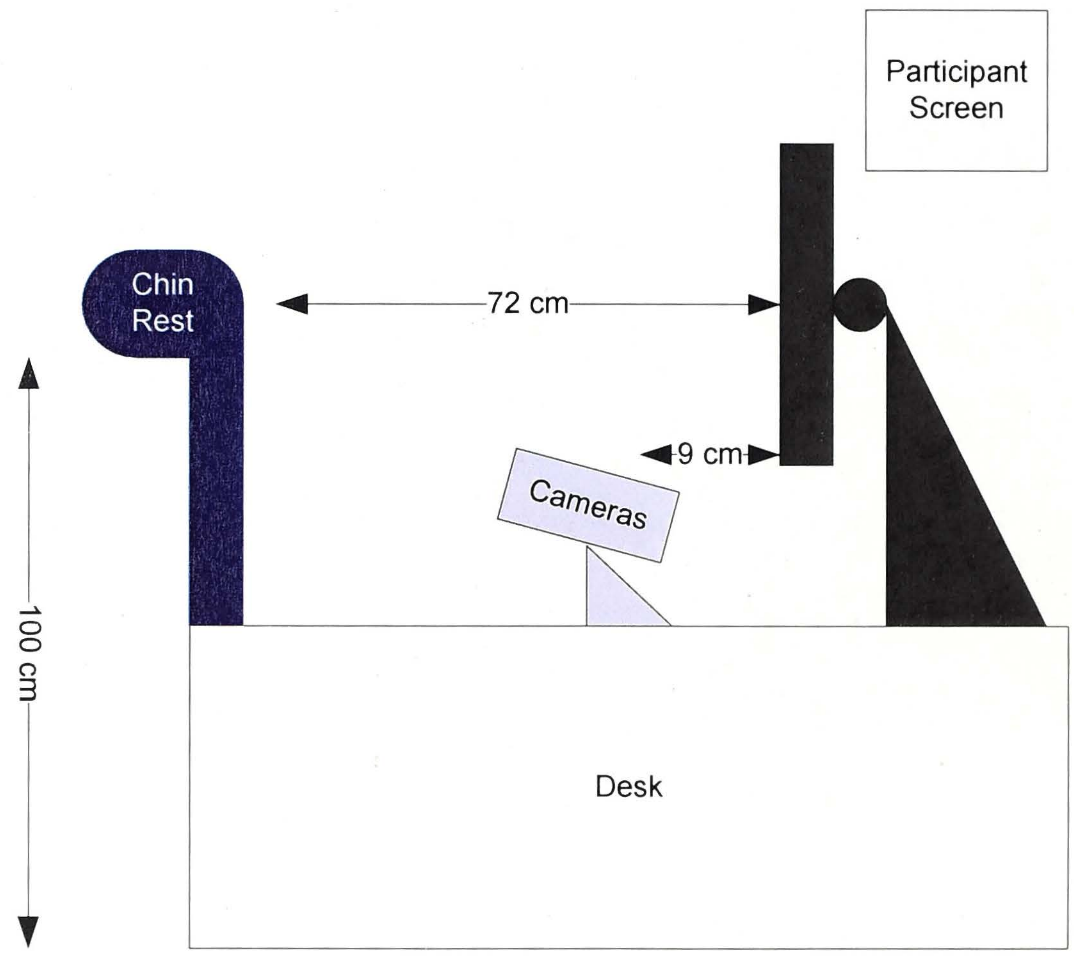


Figure 4.2: Measurements of experiment setup.

For the purpose of displaying paragraphs on screen as well as recording gaze data for *offline analysis*, a software called GazeTracker from Eye Response Technologies was used. The overall software configuration is described in Figure 4.3. The GazeTracker software has the following functions:

- It provides the facilities in scripting a screen-based and gaze-related experiment, It include the ability to *script* the order of appearance of paragraphs as well as the ability to display each paragraph on full screen by various types of triggers.

- In conjunction with the Face-Lab software, it can record the Face-lab TCP/IP packets, process them and store the data into a database.

- It also handles other matters such as timestamps, screen status, etc. All of those information is also incorporated into the same database, alongside with the gaze data.
- Data from the database can then be queried, exported for *offline analysis*.

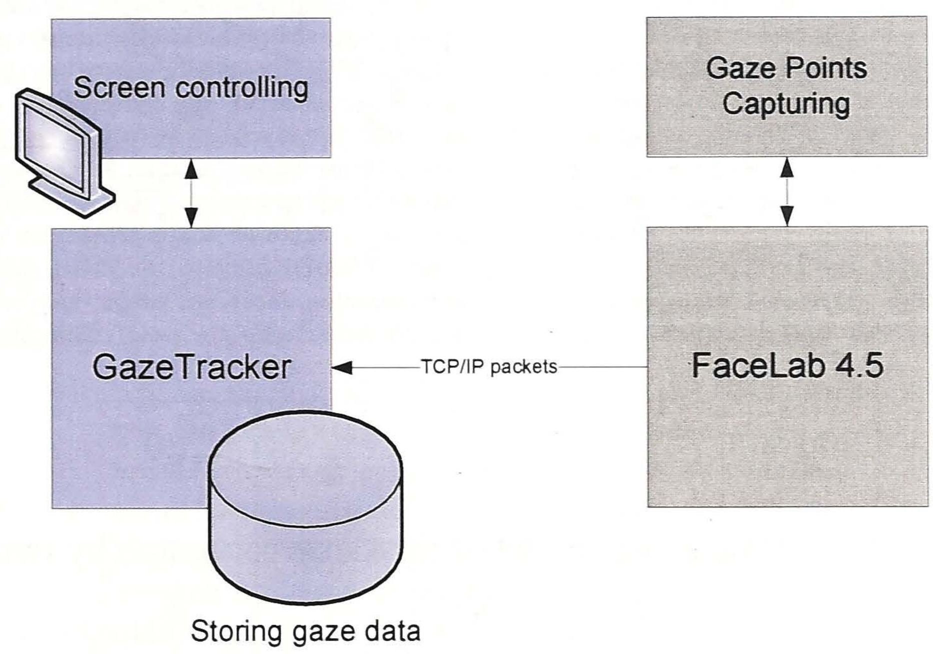


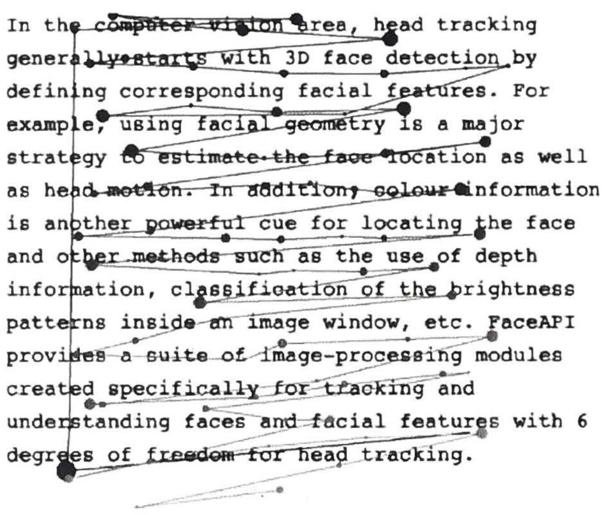
Figure 4.3: Interactions between GazeTracker and FaceLab.

4.3.5 Data Collection and Preparation

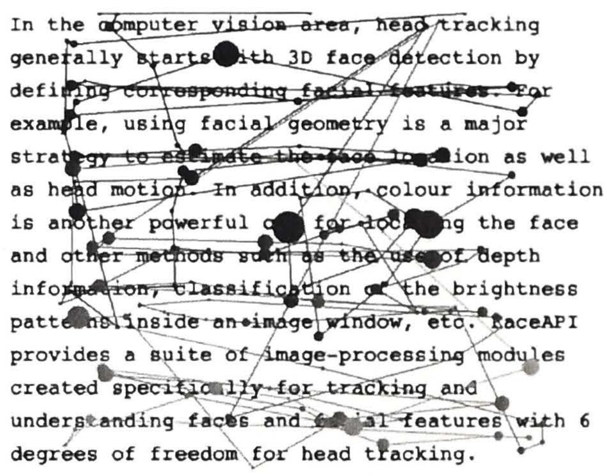
The gaze points are collected at the rate of 60 Hz and fixation points were produced from them. An approximate method was used for that purpose:

- The gaze distance threshold of 15 pixels radius [15] to define fixations was used.
- For every gaze point the distance between the previous gaze point and itself was measured (in pixels). If the distance is shorter than or equal to 15 then it is included into the cluster, or fixation, that the previous point belongs to. Otherwise, it starts a new cluster (fixation) of itself. Eventually, the gaze points that belong to fixations of themselves are discarded.
- This way, the gaze points were grouped into clusters with size of 15 pixels radius [15] to define fixations.
- The fixation length (ms) is worked out by the number of gaze points within each cluster.

Below is a visualisation of gaze data being projected onto their correspondent paragraphs. The solid circles represents the fixations. The shade of the circle indicates the fixation length - with the darker one indicate a longer fixation point. The lines that connect the fixations represent the saccades. The data was further filtered using the



(a)



(b)

Figure 4.4: An example of 2 reading patterns of the same paragraph by two different participants.

aforementioned grid-based method with 4-by-5 grid. Based on that the numbers of back-tracking and forward-tracking of each paragraph was calculated. For the classification task, these three following features to be evaluated with ANN and SVM were used:

- Average fixation length for each paragraph and each participant
- Back-tracking count for each paragraph and each participant
- Forward-tracking count for each paragraph and each participant

4.4 Proposed Method

4.4.1 Effective Reduction of Data Resolution

The gaze-tracking equipment that in used provides the gaze points in terms of a series of X and Y coordinates. These coordinates identify the locations of the gaze points on the screen and have been used to calculate the horizontal and vertical movements. In previous experiments [53][15], fixation points have been produced by filtering those gaze points, resulting in a more interpretable form for later data processing. By observation, it is found that most of the movements of fixations, i.e. saccades were just small and subtle position changes caused by the fact that the eyes did not actually focus on one place. Those saccades are considered irrelevant for this purpose and one can afford to omit them in the pattern recognition stage, hence further reduce the sample size. A simple but effective method by dividing the screen into smaller cells using a m-by-n grid was proposed. Figure 4.5 demonstrates this with a 4-by-5 grid. This effectively replaces change in positions of any two fixations with the difference in position of the cells that contain them. This is referred as **cell movements**. In the cases when the fixation movements are contained within a cell, one would consider it a no change in the cell position. The benefit of this is it will be less computationally demanding to perform any processing/analysis because the number of data points has been greatly reduced. One can also adjust the resolution of the grid (m and n) for finer or coarser filtering. The data sample from the experiment was examined with and without using

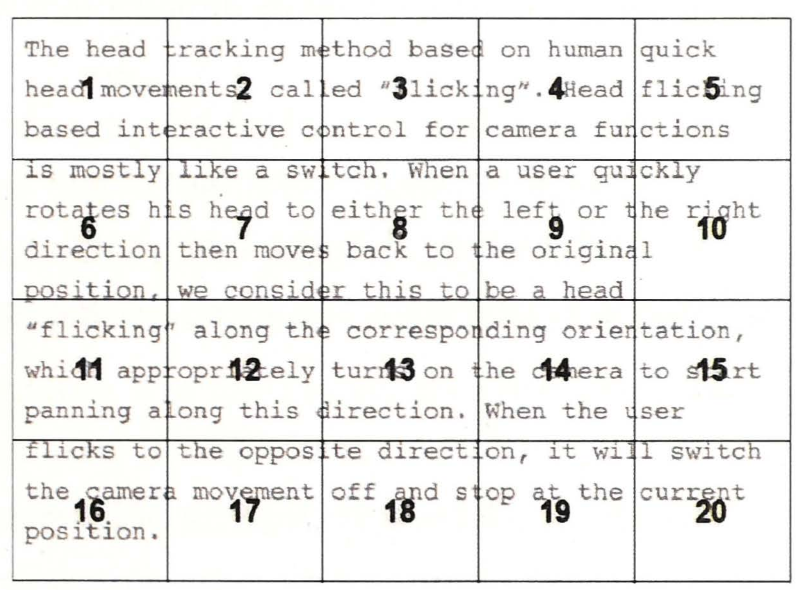


Figure 4.5: A paragraph is divided into cells by a 4-by-5 grid. Each identified by a cell number

this data reducing method. Both yielded comparable results except for computational speed.

4.4.2 Focus on Back Tracking and Forward Tracking

The most significant disruption in reading flow is the skipping forward and back-tracking activities found in the gaze. As participants try to *link* information up, they shift their eyes' focus back-and-forth to achieve a better understanding of the information. Back tracking and forward tracking are two activities that are considered as the main factor to detect engagement in reading tasks. To quantify if a gaze movement is a back/forward tracking pattern, it is considered if it belongs to the two *extreme* of cell movement groups. If one to establish a normal distribution of the distances of movement, the *extreme* groups would be the ones that did not fall within the 68-th confidence interval (a margin of one standard error). Figure 4.6 demonstrates this idea. The figure depicts the distribution of all cell movement distances of a person reading of

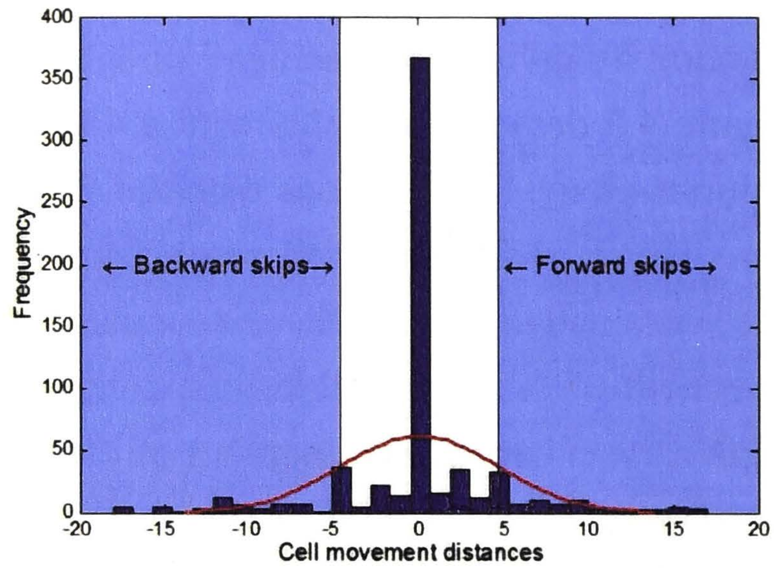


Figure 4.6: A distribution of cell movement distances throughout reading activity of a paragraph

one paragraph. It shows that if the distance of a cell movement is less than -5 ($\mu - 1\sigma$), that saccade is considered a backtracking. On the other hand, a forward tracking saccade is one that has the distance greater than 5($\mu + 1\sigma$). These thresholds are expected to be different on a case by case basis.

4.4.3 Levenberg-Marquardt based Neural Network

Previously, an experiment carried out by Zhu et al. [53] evaluating the performance of Levenberg-Marquardt optimization, combined with fuzzy signatures, in classifying gaze patterns. What they found is that this optimization algorithm performs well with the gaze-pattern classification problem and on par with SVM in the two classes test. In this section,the performance of Levenberg-Marquardt optimization as the training function in a Neural Network to classify eye gaze was evaluated. The neural network constructed is a two-layer, feed-forward back-propagation that has one single output node. As this proposed ANN model is very similar to the one described in previous chapter, please refer to Section 2.1.6 for more details.

4.5 Evaluation and Comparison

4.5.1 Neural Network Results

A two-layer neural network with one output neuron is used for classifying the data. The transfer function of the output layer is a linear transfer function while the hidden layer

Experiment	Classification Error Rate	Sensitivity	Specificity
Group A	**0.2586**	0.7241	0.7586
Group B	**0.1717**	0.7879	0.8687
Both Groups	**0.1975**	0.7898	0.8153

Table 4.1: ANN Results for Eye-gaze Feature Pattern Recognition

is equipped with a tangent sigmoid transfer function. The hidden layer comprises of 5 neurons. It is designed this to be a binary classification problem where the target values were 1 for *relevant paragraph* and 0 for *irrelevant paragraph*. The neural network was back propagation trained with Levenberg-Marquardt optimization. The LM parameters are configured with an initial μ value of 0.001 with the increase and decrease factors as 10 and 0.1 respectively. The network performance was evaluated by Mean Squared Error.

9-folds cross-validation was performed and obtained the average of classification accuracies from every fold. Due to the relatively small sample size available (35 participants - with 10 paragraphs per participant), 9-fold cross-validation is preferred to the conventional 10-fold method. For each fold, the training set is divided as followed: 60% for data training, 20% was used to generalise the network and prevent over-fitting and the last 20% was used as the test data. With this ANN setup, one were able to achieve about 80% classification accuracy as seen in Table 4.1. Figure 4.7 displays a typical performance of training the ANN with the dataset. From that chart, one can see the training process converges very quickly and the MSEs of all sub datasets (Train, Validation and Test) are close to each other consistency through out the epochs . This is encouraging because one only need to provide three gaze parameters as classification categories. As one can also observe that the classification performance achieved

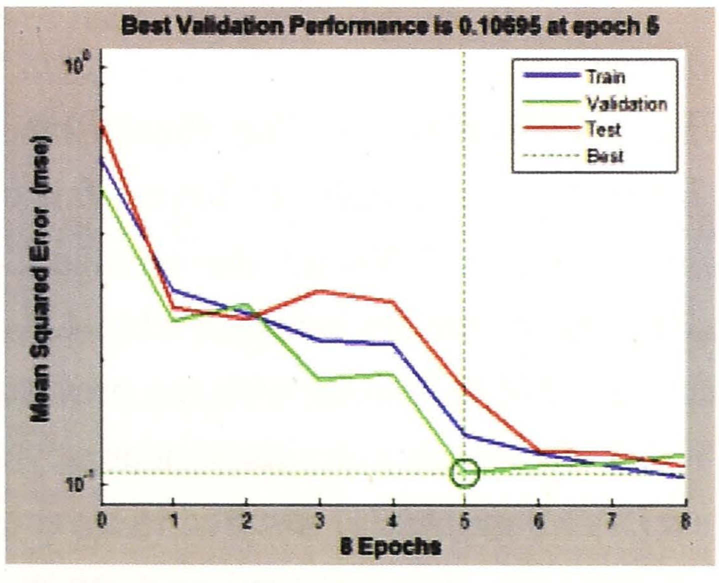


Figure 4.7: Performance (MSE) of one run of training the ANN

Experiment	SVM CER_1	ANN CER_1	SVM Sensitivity	SVM Specificity
Group A	0.254	0.259	0.7077	0.785
Group B	0.177	0.172	0.809	0.836
Both Groups	0.209	0.198	0.760	0.823

(1). Classification Error Rate

Table 4.2: SVM Results for Eye-gaze Feature Pattern Recognition

with Group A data is slightly lower than with group B data. The hypothesis behind that is that with group A, where participants were expected to answer questions about the paragraph, that lead to a more *careful* reading behaviour for all paragraphs. That would result in a less disruptive forward, backward movements in the reading patterns. But the classification results in all cases are positive considered the small number of classification features.

4.5.2 Support Vector Machine Comparison

SVM is considered a margin classifier because the purpose for its training process is to identify the maximum-margin hyperplane that can separate the training data points. Like other linear classifiers, the hyperplane is defined so that the the largest separation (hence maximum margins) between the two classes of data points is achieved. The support vectors are the data points that lie on the two margins from the hyperplane. The SVM training process is to define just that hyperplane. Please refer to Section 2.1.5 for more background information on this machine learning technique. For this classification task, the *dot product* kernel used. The kernel function k is described as follows:

$$K(x_i, x_j) = x_i^T x_j \tag{4.1}$$

The same dataset as with ANN was used. The chosen labels are "1" for *relevant paragraph* and "0" for *irrelevant paragraph*. To have a fair comparisons, the results were cross-validated using 9-folds and obtains the average Classification Error Rate (CER) after 9 iteration. The results are described in table 4.2. Table 4.2 compares the results obtained by using the SVM technique with the previous results by ANN. It is found that ANN and SVM performance in terms of accuracy almost exactly matched each other. Both methods (SVM and ANN) resulted in a very high accuracy rate and with further optimisations on both, it is believed one can attain even more positive results.

4.5.3 Comparison with EEG Results

In this section, the result of gaze-tracking data were also compare with the results achieved with EEG from the previous experiment. The comparison is made with the result of Group B dataset where the experiment set-up is almost identical to the previous experiment. The results are shown in Table 4.3. One can see that the overall classification accuracy obtain with gaze-tracking is almost on par with the one obtained from studying EEG data. The difference in values among the three parameters is very negligible, suggesting that the correlation between the two measuring techniques is valid.

Table 4.3: ANN classification results for 19 participants (together)

Method	Accuracy	Specificity	Sensitivity
EEG - ANN	**0.817**	0.889	0.716
Gazetrack	**0.828**	0.869	0.788

4.6 Future work and Opportunities

4.6.1 Analysing other information gathered from participants of Group A

The analysis and study of the following data were not part of this chapter:

- The answers to the questions of each participant in part 2.
- The gaze data of the duration when the participants attempt to find the most appropriate answer for each question.
- The answers to the questions of each participant in part 3.
- The rating of paragraphs in term of *relevance* from each participant in the group.

In general, the info used in part 2 and part 3 of this group was not analysed. Was this because these information was not helpful at all? The only reason for them not being analysed in this work is because they are out of my original scope. Fahey [15], who attempted to statistically analyse these information had indicated that "Even though the statistical trends against the scores are weak, the data from the experiment is still useful."[15]. There is, however, a need for a valid and decent *hypothesis* that could be made from the information above. With the recent advances in statistical machine classifiers, the statement of Fahey's would imply that there is always a potential in using machine learning techniques to validate and confirm it.

4.6.2 Attempt at different sets of feature

These are the list of features that have the potential to be included into the feature set for the machine learning classifiers:

- Average time taken to read a slide
- Average number of gaze points on a slide
- Average number of fixations on a slide
- Average duration of a fixation
- Average horizontal movement between fixations
- Average vertical movement between fixations

The reason those were not used as part of this work is because the chosen feature set is already considered *good enough.* A further investigation on the use of these statistical attribute is recommended. However, the advantage of these features is that they are easier to compute than the current feature set.

4.7 Summary

In this chapter, the effectiveness of ANN in recognising gaze-patterns was demonstrated. The findings are encouraging because ANN combined with the proposed method for data pre-processing has resulted in a low computational cost model that achieves highly accurate results: An ANN trained with only three features dataset is able to achieve 80% accuracy is very encouraging. It has also consolidated the outcome of previous experiment [15] as well as the use of Levenberg-Marquardt optimization as the training algorithm for this types of problem [53]. Moreover, the results of this study, from the preliminary analysis stage throughout the final results have shown that there are certain relationships between EEG signals captured from the human brain to the way a person reads or perceives the information while reading. Section 5.1.3 further emphasis this point.

This work has been published into the proceeding of *International Conference on Neural Information Processing 2010* conference organised in Sydney, 2010. [51].

Chapter 5

Conclusions and Future Work

5.1 Conclusion

5.1.1 Explicit Differentiation and EEG

The conclusion starts with looking back to the results achieved from **Chapter 2**. The trials have demonstrated *very good* potentials in the ability to identify the cognitive process of *explicit differentiation* from studying the EEG signals. The use of *machine learning* tools that achieve high correct rate of 80 percent classification results in most cases has further strengthened the belief that this is the right track to be able to achieve that goal, at least in *statistical* terms. Provided a more significant number of test subjects would likely to further consolidate that belief.

The chapter's Discussion section expressed some concerns, especially, on the original experiment design. But overall, the results achieved are far better from original expectation. It is encouraging to see good results obtained through the use of EEG. I highly recommend this preliminary study to be followed up by proper trials and more sizeable number of test subjects. That way, one can properly *confirm* the validity of the claims in the chapter, setting ourselves to be in a very good position of constructing a working *online* BCI system as originally planned.

5.1.2 Implicit Differentiation and EEG

Chapter 3 is dedicated to the investigation of *implicit differentiation* by studying the EEG signals captured during reading tasks. The positive classification result and a significantly sized dataset have confirmed the hypothesis that with EEG, one can effectively detect differentiating activities occurring during a reading task. This chapter also demonstrates the two following points:

- It consolidates the findings from **Chapter 2**. By using similar methods of EEG signal processing and classification tools, this result is an extension to support the claim that it is possible to computationally detect *various* differentiation types by studying EEG signals.
- It demonstrates the potential for an online BCI system for the purpose, thanks to the use of effective and light-weight signal processing methods in the experiment set-up.

5.1.3 EEG and Eye-gaze Correlations

As for the study of reading tasks in **Chapter 3** and **Chapter 4**, one can *conclude* that using either *EEG* or *gaze-tracking* would yield very *similar results*. This is in many way encouraging because EEG performs *very well* compared to a measuring method that is *specifically* designed for the task. It also shows that, EEG, normally being associated with *noisy signals* and its *stochastic nature*, has good potentials to be considered a proper BCI input mechanism.

In my opinion, it is more important to put this result in another perspective: there exists some correlation between the classification performances of gaze-tracking and EEG data. There is a potential in using gaze-tracking together with EEG to further improve the spatial resolution of EEG.

5.2 Future Work

5.2.1 Signal Processing

General

One of the comments received from the first experiment with EEG is that to deal with *noisy* signals like EEG, the usual approach of signal processing was not effective enough. (Filtering and FFT etc.).

There are extensive range of signal processing tools that could be used in this area that I have not had the opportunity to properly explore during my time here: ICA, BSS or simply other types of filters (static / adaptive) that may raised better results.

Feature Extraction

There are quite a few ideas that I could use to improve the current results:

- Dealing with imbalanced data sets: I did run into over-fitting problem with the dataset I obtained from the first experiment. It is one of the indications that I am dealing with imbalanced datasets.

- Feature extractions with tensor-based techniques have been used in RIKEN Brain Science Institute for diagnosing with some success in diagnosing complex brain symptoms such as epilepsy.

5.2.2 Investigate other Machine Learning Methods

Alternative SVM Kernel Methods

A new method is the use of sphere-based/multisphere-based kernel methods. Trung [34] has commented that the performance of these kernels used for EEG signals are on par, if not better than the existing SVN kernels. The evidences provided by Trung [34] seem to confirm his comments and could benefit this work in the future.

Nonnegative Tensor Factorization for Classification

NTF is also considered another alternative for SVM as mentioned from the feature extraction section 5.2.1. However, more time and effort needed to explore this potential research work, which has been demonstrate by Lee et al. [36].

Appendix A

Reading Paragraphs

In the computer vision area, head tracking generally starts with 3D face detection by defining corresponding facial features. For example, using facial geometry is a major strategy to estimate the face location as well as head motion. In addition, colour information is another powerful cue for locating the face and other methods such as the use of depth information, classification of the brightness patterns inside an image window, etc. FaceAPI provides a suite of image-processing modules created specifically for tracking and understanding faces and facial features with 6 degrees of freedom for head tracking.

(a) Paragraph 1 (Relevant)

The head tracking method, called "motion", operates according to natural human head motion. Assuming initially that the user's head is directly facing the screen, when the user rotates the head to either left or right by a certain angle, the camera will pan in the corresponding direction. It will keep panning the view along that direction until the user moves their head back to the original position. When the user tilts their head up or down by a certain angle, the camera will correspondingly carry out the tilt function and not stop tilting until the head returns to the original position.

(b) Paragraph 2 (Relevant)

The head tracking method based on human quick head movements, called "flicking". Head flicking based interactive control for camera functions is mostly like a switch. When a user quickly rotates his head to either the left or the right direction then moves back to the original position, we consider this to be a head "flicking" along the corresponding orientation, which appropriately turns on the camera to start panning along this direction. When the user flicks to the opposite direction, it will switch the camera movement off and stop at the current position.

(c) Paragraph 3 (Relevant)

Our objective results indicate that for this specific experimental setting, keyboards still performed the best by most of the subjects. We believe this is due to the fact that all the participants were quite familiar with using the keyboard, and initially there was no training time for them to get used to the two head tracking control methods. The reason for requiring the subjects to immediately start performing the experiment was to test how well users could pick up the head tracking based remote control. It is clear that our "head motion" based design provides quite comparable performance to the most conventional device (keyboard) even without any training.

(d) Paragraph 4 (Relevant)

In this research, we focus on importing computer vision technology to undertake head tracking in interface design for teleoperation activities. The common remote control situation described above is modelled by using a physical game analogue: playing a table soccer game with two handles. This has the advantage of being more compelling for our student experimental subjects than a more abstract task. We use student experimental subjects as we have limited access to the operators. We then propose a novel design applying natural human head gestures for controlling a Pan-Tilt-Zoom camera as an effective approach to solve the camera control problem.

(e) Paragraph 5 (Relevant)

Head tracking is a key component in applications such as human computer interaction, person monitoring, driver monitoring, video conferencing, and object-based compression. Recently, one of the most popular ways of applying head tracking is to couple the virtual camera to a user's head position in order to achieve a more realistic and immersive experience of perspective in virtual reality or visual gaming.

(f) Paragraph 6 (Irrelevant)

There have also been attempts to develop head tracking based "hands-free pointing" interface for controlling the mouse cursor, by which a user can point their nose where they wish to place the cursor on a monitor screen. "hMouse" is another head tracking driven camera mouse system, which provides alternative solutions for convenient device control with potential applications for people with disabilities and the elderly.

(g) Paragraph 7 (Irrelevant)

Technology is constantly changing and evolving around us. And, as technology evolves, so does the media that is viewed and accessed through new mediums. New technologies allow for greater flexibility and more accessibility. Such technologies are now rising that pose the threat of making past technologies obsolete.

(h) Paragraph 8 (Irrelevant)

Participants in a research experiment used various sensory methods in order to find pictures of human faces and note the numbers on them with a remote camera view. Participants were then asked to write a research report on their experiences and how or whether they relate to their studies of the World Wide Web. Human-computer interaction (HCI) is always pressing forward into the future, and this experiment is one of many that will help it do so in the years to come.

(i) Paragraph 9 (Irrelevant)

The accuracy and complexity of human vision and movement is a concept which has been studied for years. Extensive study and research has been conducted through out history to create human computer interface designs such has head-tracking technology, joysticks and other control methods of machines to help with (in particular) labour-intensive industries. However technology is constantly striving to imitate natural human movement, which means research and study, is continuously being conducted to help improve and test new designs.

(j) Paragraph 10 (Irrelevant)

Bibliography

[1] Researchers use brain interface to post to Twitter. Website, April 2009. `http://www.news.wisc.edu/16576`.

[2] EEGLAB - Open Source Matlab Toolbox for Electrophysiological Research. Website, September 2010. `http://sccn.ucsd.edu/eeglab/`.

[3] Biosemi EEG ECG EMG BSPM NEURO amplifier electrodes. Website, September 2011. `http://www.biosemi.com/products.htm`.

[4] Emotiv — EEG System — Electroencephalography. Website, September 2011. `http://http://emotiv.com/`.

[5] Features. Website, September 2011. `http://www.bci2000.org/BCI2000/Features.html`.

[6] Tan Le: A headset that reads your brainwaves — Video on TED.com. Website, October 2011. `http://www.ted.com/talks/tan_le_a_headset_that_reads_your_brainwaves.html`.

[7] M. Ali and J. Ole. Posterior alpha activity is not phase-reset by visual stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 103(8):2948–2952, February 2006.

[8] J. Bonnie and Kaplan. Morphological evidence that feline SMR and human Mu are analogous rhythms. *Brain Research Bulletin*, 4(3):431–433, 1979.

[9] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM Press, 1992.

[10] E. Byvatov, U. Fechner, J. Sadowski, and G. Schneider. Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Clas-

sification. *Journal of Chemical Information and Computer Sciences*, 43(6):1882–1889, 2003.

[11] A. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. S.Raizada. NeuroPhone: brain-mobile phone interface using a wireless EEG headset. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, MobiHeld'10, pages 3–8, New York, NY, USA, 2010. ACM Press.

[12] I. A. Cook, R. O'Hara, S. H. Uijtdehaage, M. Mandelkern, and A. F. Leuchter. Assessing the accuracy of topographic EEG mapping for determining local brain function. *Electroencephalography and Clinical Neurophysiology*, 107(6):408–414, 1998.

[13] J. W. Cooley and J. W. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, 19(90):297–301, 1965.

[14] D. A. Craig and H. T. Nguyen. Adaptive EEG Thought Pattern Classifier for Advanced Wheelchair Control. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2544–2547, August 2007.

[15] F. Daniel. A Preliminary Investigation into using Eye-Tracking to Analyse a Persons Reading Behaviour. Master's thesis, The School of Computer Science Australian National University, 2009.

[16] C. Dietrich. Decision Making: Factors that Influence Decision Making, Heuristics Used, and Decision Outcomes. *Student Pulse*, 2(2), 2010.

[17] E. Edwards, M. Soltani, L. Y. Deouell, M. S. Berger, and R. T. Knight. High Gamma Activity in Response to Deviant Auditory Stimuli Recorded Directly From Human Cortex. *Journal of Neurophysiology*, 94(6):4269–4280, 2005.

[18] E. Eweda. Signal behavior of adaptive filtering algorithms in a nonstationary environment with singular data covariance matrix. *Signal Processing*, 85(6):1263–1274, 2005.

[19] A. Ferreira, R. L. Silva, W. C. Celeste, T. F. B. Filho, and M. S. Filho. Human machine interface based on muscular and brain signals applied to a robotic wheelchair. *Journal of Physics Conference Series*, 90(1):1541–1672, 2007.

[20] M. Frigo and S. G. Johnson. The Design and Implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, February 2005.

[21] F. Galn, M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, and J. del R. Milln. A brain-actuated wheelchair: asynchronous and non-invasive Brain-computer interfaces for continuous control of robots. *Clinical Neurophysiology*, 119(9):2159–2169, 2008.

[22] M. George. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, 63:81–97, 1956.

[23] A. B. A. Graf, A. J. Smola, and S. Borer. Classification in a normalized feature space using support vector machines. *Neural Networks, IEEE Transactions on*, 14(3):597–605, May 2003.

[24] M. T. Hagan and M. B. Menhaj. Training feedforward networks with the Marquardt algorithm. *Neural Networks, IEEE Transactions on*, 5(6):989–993, November 1994.

[25] H. Hallez, A. Vergult, R. Phlypo, P. V. Hese, W. D. Clercq, Y. DAsseler, R. V. D. Walle, B. Vanrumste, W. V. Paesschen, and S. V. Huffel. Muscle and eye movement artifact removal prior to EEG source localization. *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1(2):1002–1005, 2006.

[26] S. Hanslmayr, W. Klimesch, P. Sauseng, W. Gruber, M. Doppelmayr, R. Freunberger, T. Pecherstorfer, and N. Birbaumer. Alpha Phase Reset Contributes to the Generation of ERPs. *Cerebral Cortex*, 17(1):1–8, 2007.

[27] C. J. Henry. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Fifth Edition. *Neurology*, 67(11):2092–2092–a, 2006.

[28] A. R. Hobson and A. Hillebrand. Independent component analysis of the EEG: is this the way forward for understanding abnormalities of braingut signalling? *Gut*, 55(5):597–600, 2006.

[29] C. W. Hsu, C. C. Chang, and C. J. Lin. A Practical Guide to Support Vector Classification. Technical report, Department of Computer Science, National Taiwan University, 2003.

[30] S. S. Jonathan. The local mean decomposition and its application to EEG perception data. *Journal of The Royal Society Interface*, 2(5):443–454, December 2005.

[31] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME Journal of Basic Engineering*, (82 (Series D)):35–45, 1960.

[32] Y. Kikuchi, M. Yoshii, K. Yanashima, T. Enoki, T. Ide, F. Sakemi, and S. Okisaka. Multifocal Visual Evoked Potential Is Dependent on Electrode Position. *Japanese Journal of Ophthalmology*, 46(5):533–539, 2002.

[33] A. I. Klistorner, S. L. Graham, J. R. Grigg, and F. A. Billson. Multifocal topographic visual evoked potential: improving objective detection of local visual field defects. *Investigative Ophthalmology and Visual Science*, 39(6):937–50, 1998.

[34] T. Le, D. Tran, W. Ma, and D. Sharma. A Theoretical Framework for Multisphere Support Vector Data Description. In Kok Wong, B. Mendis, and Abdesselam Bouzerdoum, editors, *Neural Information Processing. Models and Applications*, volume 6444 of *Lecture Notes in Computer Science*, pages 132–142. Springer Berlin / Heidelberg, 2010.

[35] M. A. Lebedev and M. A. Nicolelis. Brainmachine interfaces: past, present and future. *Trends in Neurosciences*, 29(9):536 – 546, 2006.

[36] H. Lee, Y. Kim, A. Cichocki, and S. Choi. Nonnegative tensor factorization for continuous EEG classification. *International Journal of Neural Systems*, 17(4):305–317, 2007.

[37] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of neural engineering*, 4(2), June 2007.

[38] I. S. Macleod, G. Hone, and S. Smith. *Capturing Cognitive Task Activities for Decision Making and Analysis*, volume 44, pages 0–20. 2005.

[39] S. Makeig, S. Debener, J. Onton, and A. Delorme. Mining event-related brain dynamics. *Trends in Cognitive Sciences*, 8(5):204–210, 2004.

[40] S. Makeig, M. Westerfield, T. P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. J. Sejnowski. Dynamic Brain Sources of Visual Evoked Responses. *Science*, 295(5555):690–694, 2002.

[41] O. Mariko. Peak Alpha Frequency of EEG during a Mental Task: Task Difficulty and Hemispheric Differences. *Psychophysiology*, 21(1):101–105, 1984.

[42] B. S. U. Mendis, T. D. Gedeon, and L.T. Koczy. Learning Generalized Weighted Relevance Aggregation Operators Using Levenberg-Marquardt Method. In *Hybrid Intelligent Systems, 2006. HIS '06. Sixth International Conference on*, page 34, December 2006.

[43] C. Neuper and G. Pfurtscheller. Evidence for distinct beta resonance frequencies in human EEG related to specific sensorimotor cortical areas. *Clinical Neurophysiology*, 112(11):2084 – 2097, 2001.

[44] J. V. Odom, M. Bach, C. Barber, M. Brigell, M. F. Marmor, A. P. Tormene, G. E. Holder, and Vaegan. Visual evoked potentials standard (2004). *Documenta ophthalmologica. Advances in ophthalmology*, 108(2):115–123, March 2004.

[45] S. H. Patel and P. N. Azzam. Characterization of N200 and P300: selected studies of the Event-Related Potential. *International journal of medical sciences*, 2(4):147–154, January 2005.

[46] B. Rebsamen, E. Burdet, C. Guan, H. Zhang, C. L. Teo, Q. Zeng, C. Laugier, and M. H. Ang Jr. Controlling a Wheelchair Indoors Using Thought. *Intelligent Systems, IEEE*, 22(2):18–24, March 2007.

[47] J. M. Schraagen, S. F. Chipman, and V. L. Shalin. *Cognitive Task Analysis*, volume 25. L. Erlbaum Associates, 2000.

[48] M. M. Shaker. EEG Waves Classifier using Wavelet Transform and Fourier Transform. *International Journal of Biological Biomedical and Medical Sciences*, 1(2):85–90, 2006.

[49] K. Srivastava and V. P. Dimri. Adaptive filter for processing of multichannel nonstationary seismic data. In *Multidimensional Signal Processing Workshop, 1989., Sixth*, page 50, September 1989.

[50] R. VanRullen and S. J. Thorpe. The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4):454–61, 2001.

[51] T. Vo, B. S. U. Mendis, and T. D. Gedeon. Gaze Pattern and Reading Comprehension. In Kok Wong, B. Mendis, and Abdesselam Bouzerdoum, editors, *Neural Information Processing. Models and Applications*, Lecture Notes in Computer Science, pages 124–131. Springer Berlin / Heidelberg, 2010.

[52] D. Zhu, T. D. Gedeon, and K. Taylor. Keyboard before Head Tracking Depresses User Success in Remote Camera Control. In *INTERACT (2)*, pages 319–331, 2009.

[53] D. Zhu, B. S. U. Mendis, T. D. Gedeon, A. Asthana, and R. Goecke. A Hybrid Fuzzy Approach for Human Eye Gaze Pattern Recognition. In *ICONIP (2)'08*, pages 655–662, 2008.